MONTHLY NOTICES

OF THE

ROYAL ASTRONOMICAL SOCIETY

Vol. 116 No. 6 1956

Published and Sold by the

ROYAL ASTRONOMICAL SOCIETY

BURLINGTON HOUSE

LONDON, W.1

Price Thirteen Shillings and Sixpence

ROYAL ASTRONOMICAL SOCIETY

Founded 1820

OCCASIONAL NOTES

of the

ROYAL ASTRONOMICAL SOCIETY

No. 19 of Volume 3 consists of a single article:

Close Binary Stars. By OTTO STRUVE and SU SHU HUANG

Now available: price 4s. 6d.; in U.S.A. $0.80.

MEMOIRS

of the

ROYAL ASTRONOMICAL SOCIETY

Volume LXVII, published during 1954/55, consists of four parts. These are on sale individually at the following prices:

Part I. **Photographic and photovisual magnitudes of 7^m–10^m stars in the $+15°$ Selected Areas.** By A. BEER, R. O. REDMAN and G. G. YATES.
Price £1 0s. 0d.; in U.S.A. $3.20.

Part II. **Radial velocities of southern B stars determined at the Radcliffe Observatory (Paper I).** By M. W. FEAST, A. D. THACKERAY and A. J. WESSELINK.
Price £1 2s. 0d.; in U.S.A. $3.50.

Part III. **The Cambridge radio telescope.** By M. RYLE and A. HEWISH.
A survey of radio sources between declinations $-38°$ and $+83°$. By J. R. SHAKESHAFT, M. RYLE, J. E. BALDWIN, B. ELSMORE and J. H. THOMSON.
Price £1 8s. 0d.; in U.S.A. $4.30.

Part IV. **A survey of southern H II regions.** By COLIN S. GUM.
Price £1 12s. 0d.; in U.S.A. $4.90.

Orders for the above should be addressed to:

THE ASSISTANT SECRETARY,
Royal Astronomical Society, Burlington House, London, W.1

MONTHLY NOTICES

OF THE

ROYAL ASTRONOMICAL SOCIETY

Vol. 116 No. 6

MEETING OF 1956 NOVEMBER 9

Professor Sir Harold Jeffreys, President, in the Chair

The election by the Council of the following Fellows was duly confirmed :—

Alan Kingsley Appleton, 4 Beechwood Avenue, Chorley Wood, Herts. (proposed by N. J. Goodman);

Fred T. Haddock, The Observatory, University of Michigan, Ann Arbor, Michigan, U.S.A. (proposed by N. G. Roman);

Norman S. Jinkinson, 81 Meersbrook Park Road, Sheffield 8 (proposed by R. R. S. Cox);

Peter Hendrikus Mans, Ruyterstraat 13, Middelharnis (Z–H), Holland (proposed by G. Tait);

VR. Ramanathan, Royal Holloway College, Englefield Green, Surrey (proposed by W. H. McCrea);

Percival Albert Sheppard, Imperial College, London, S.W.7 (proposed by G. E. R. Deacon);

Leonard Stuart Strickson, 26 Ripon Road, Harrogate, Yorkshire (proposed by G. Fielder); and

Raymond Thomas, 89 Factory Road, Hinckley, Leics. (proposed by H. Bondi).

The election by the Council of the following Junior Members was duly confirmed:—

Michael George Julian Van der Burg, 7 Eatonville Road, Tooting, London, S.W.17 (proposed by H. Bondi); and

Derrick Alan Watt, 105 Grove Park Road, Chiswick, London, W.4 (proposed by H. Bondi).

Seventy-three presents were announced as having been received since the last Meeting, including:—

A. Vibart Douglas, *The Life of Arthur Stanley Eddington* (presented by the author);

A. Dauvillier, *L'origine des planètes*; and

A. Dauvillier, *Cosmologie et chimie* (both presented by the publishers, Presses Universitaires de France).

MEETING OF 1956 DECEMBER 14

Professor Sir Harold Jeffreys, President, in the Chair

The election by the Council of the following Fellows was duly confirmed :—

Rev. Canon Daniel John Amies, The Vicarage, Marlow, Bucks. (proposed by C. T. Daltry);

Graham Ashworth, Beckington Abbey, Beckington, Bath, Somerset (proposed by J. C. Rome);

John David Bestwick, Iddesleigh, Hartington Street, Leek, Staffs. (proposed by P. A. Moore);

Michael A. Browne, 24a Oxford Road, Manchester 1 (proposed by E. Burgess);

David Cremer Cooke, 9 Forest Court, Snaresbrook, London, E.11 (proposed by C. W. Allen);

Charles Arthur Cross, 284 London Road, Northwich, Cheshire (proposed by P. A. Moore);

Diana Elizabeth Damen, Royal Greenwich Observatory, Abinger Common, Dorking, Surrey (proposed by R. v. d. R. Woolley);

Joseph Clark Davidson, 12 Kersland Street, Glasgow, W.2 (proposed by W. M. Smart);

Olin Jeuck Eggen, Royal Greenwich Observatory, Herstmonceux Castle, Hailsham, Sussex (proposed by R. v. d. R. Woolley);

John Mortleman, 130 Pampisford Road, Purley, Surrey (proposed by F. J. Hargreaves);

Nathy Patrick Joseph O'Hora, Royal Greenwich Observatory, Abinger Common, Dorking, Surrey (proposed by R. v. d. R. Woolley);

Brian William Petley, Lyndale, Old Hill, Egley Road, Woking, Surrey (proposed by Z. Kopal);

Nadaratah Shanmukharatnam, William Temple House, 29-31 Trebovir Road, London, S.W.5 (proposed by W. F. Kibble);

David Sinden, 6 Haddon Road, Billingham-on-Tees, Co. Durham (proposed by H. P. Wilkins);

Theodore Eugene Sterne, Smithsonian-Harvard College Observatories, 60 Garden Street, Cambridge 38, Mass., U.S.A. (proposed by C. Payne Gaposchkin); and

Saadettin Topuzoglu, Besiktas, Misirli Bahce Sokak No. 99, Istanbul, Turkey (proposed by H. P. Wilkins).

The election by the Council of the following Junior Member was duly confirmed:—

Derek Hugh Powell Jones, Christ's College, Cambridge (proposed by R. v. d. R. Woolley).

One hundred and one presents were announced as having been received since the last Meeting, including:—

Academy of Sciences, U.S.S.R., *Discussions on Cosmogony*, Vols. II, III and IV (in Russian);

A. Beevar, *Atlas Coeli* 1950.0 (presented by the Czech Academy of Sciences);

E. A. Beet, *Teaching astronomy in schools* (presented by the author); and

H. M. Nautical Almanac Office, *Interpolation and allied tables*.

CHROMOSPHERIC EMISSION AND ITS DEPENDENCE ON LATITUDE

R. O. Bishop

(Communicated by the Director, University Observatory, Oxford)

(Received 1956 October 1)

Summary

The observation and measurement of the total intensities at a succession of heights in the chromosphere of the emission lines, hydrogen Hα and helium D_3, at three series of latitudes on the limb of the Sun, are described. Results expressed in absolute units by comparison with the photosphere are tabulated for heights 2700 km and 4500 km in the chromosphere for the latitudes of the the three series. Investigation of the latitude dependence of the emission indicates a general trend with greater emission at the poles than at the equator for Hα and the reverse for D_3, together with local irregularities. An attempt has been made to correct the observed height distribution of emission for the effects of seeing.

Introduction.—Properly-timed spectroscopic observations at an eclipse have the over-whelming advantage of showing the transition zone between the photosphere and the chromosphere. They suffer, however, from the disadvantages of being confined in practice to low solar latitudes, of not readily yielding absolute surface brightness, and of being limited to short irregularly-spaced time-intervals. Observations outside an eclipse, on the other hand, have their own peculiar difficulties, but at least the emission of a line can be found as a function of latitude and expressed in absolute units by a direct comparison with the continuous spectrum of photospheric radiation. Unsöld (**17**) (**18**), Keenan (**7**), Perepelkin and Melnikov (**12**), and de Jager (**6**) have made quantitative chromospheric observations outside an eclipse and have shown how to overcome some of the difficulties of this type of observation, but no quantitative investigation has yet been made of the dependence, if any, of chromospheric emission on latitude. It is the purpose of the present paper to present a preliminary investigation of this dependence for the hydrogen line Hα and the helium line D_3. The first two sections describe the methods of observation, measurement and reduction ending with a table of results. The third section discusses the variation with latitude and the final section the results of an attempt to find the variation of the emission with radial height in the chromosphere.

1. *Observation.*—To find the intensities of chromospheric lines outside an eclipse it is desirable to have a large solar image. This requirement is satisfactorily met by the 35·2 m telescope (**15**) with a scale of approximately 5·9 seconds of arc per mm, so that with a radial slit the chromosphere has a linear height of some 1·7 mm. Since the large spectroscope to be used with this telescope is not yet installed, a Hilger Medium Glass Spectrograph was mounted with its slit in the focal plane of the telescope. The dispersion of this spectroscope is

65·1 A mm^{-1} at Hα and 45·3 A mm^{-1} at D_3. Because of the large tilt of its focal plane the linear magnification of this spectroscope is such that 1 mm on the slit corresponds to a height of spectrum of 0·972 mm at Hα and 0·957 mm at D_3.

With a radial slit crossing the limb of the Sun and covering the chromosphere, a large amount of bright photospheric light gets into the spectroscope. To reduce scattering in the instrument an Ilford No 202 filter, cemented between optical flats and with a sharp cut-off to the violet of 5600 A, was mounted in front of the slit. A Compur shutter of 3 cm clear aperture, well behind the slit, formed a light-tight trap in front of the collimator, and allowed "instantaneous" exposures of precisely equal duration to be made on each spectrum. Ilford R.40 Rapid Process Panchromatic plates were used throughout and developed with Parkhurst developer, normally by the brush method and after storage for 24 hours.

The solar image was adjusted so that the spectroscope slit cut radially across the upper limb. Because the image at the Coudé focus rotates uniformly the position angle, measured from north through east, of the slit is given by

$$\phi = 231^{\circ}\cdot 9 - 15h$$

where h is the hour angle of the Sun at Oxford. The angle, ψ, at the centre of the solar disk between the north pole of rotation of the Sun and the spectroscope slit is

$$\psi = \phi - P$$

where P is the position angle of the north pole. The solar latitude, B, of the slit, given by

$$\sin B = \sin S \sin B_0 + \cos S \cos B_0 \cos \psi$$

where S is the angular semi-diameter and B_0 the latitude of the centre of the disk, changes steadily throughout the day. P, B_0, and S are given for each day in the Nautical Almanac.

On 1954 August 27, a day free from cloud but slightly hazy, 14 plates were obtained by H. H. Plaskett with this observational arrangement. The latitude of the slit changed during the period of observation from -67° to -82° and from -82° to $+5^{\circ}$. The visible disk of the Sun was completely free of spots. Every plate carried six spectra, each with an exposure of 0·2 seconds. Three of these were of the limb of the Sun and chromosphere taken in the way already described and three were of the solar photosphere. One of these latter was a spectrum of 1 mm height taken a known distance inside the limb, and two were spectra of the centre of the solar image. Of these centre spectra one was taken through a Hilger step wedge, described and calibrated by R. J. Bray (**1**), and the other through the blank glass of the wedge adjacent to the rhodium steps. In order to keep the blackening of these three photospheric spectra comparable with that of the chromospheric spectra, neutral tint filters, calibrated by the National Physical Laboratory, were placed in front of the slit in addition to the Ilford No. 202 filter. The wedge spectrum of the centre of the Sun's disk, corrected by the "blank" spectrum for any possible non-uniformity of illumination, gives the plate calibration curve.

These spectra, referred to hereafter as Series I, were not entirely satisfactory. From them, as described in the next section, the total intensities of Hα and D_3 can be found in erg cm^{-2} sec^{-1} per unit solid angle from the spectra of the

photosphere. Three such photospheric spectra are available on each plate, but none was exposed simultaneously with a spectrum of the chromosphere. Any variation of atmospheric transmission between these exposures therefore leads to an error in the finally determined total intensity. Though the sequence of exposures was so arranged that any secular change of transmission led only to an accidental error, the results from Series I showed an average r.m.s. error for a single determination of 10 per cent from the three photospheric spectra. Another unsatisfactory feature of this series of spectra was the over-exposure of the solar limb in the chromospheric spectra. This made it difficult to determine the intensity distribution across the limb, a quantity required both to fix the true position of the limb and find the radius of the tremor disk.

To overcome these difficulties, I suggested that the chromospheric spectra should be exposed through the Hilger step wedge in such a way that the bright photospheric limb light was transmitted only by the denser steps of the wedge. With this improvement, which provides simultaneous exposure of photospheric and chromospheric spectra, but with no other change of technique two series of spectra (Series II and III) were obtained by the same observer in 1955. On each plate four spectra were exposed, two independent spectra each showing the photospheric limb and the chromosphere exposed through the wedge, and two calibrating spectra of the centre of the Sun taken through the steps and blank glass of the wedge as in the first series. Series II consists of 12 such plates taken on 1955 May 30. The exposure times for all spectra on any one plate are identical, but varied from plate to plate between 0·35 and 0·50 seconds. The latitude of the slit varied from $-78°$ to $-23°$ and from $+15°$ to $+50°$. Series III consists of another 12 plates taken on 1955 May 31 with identical exposure times on each plate, but varying from plate to plate between 0·20 and 0·50 seconds. In this series the latitude of the slit varied from $-3°$ to $+65°$. On both days the sky was brilliantly clear, and only two spots (less than 5″ of arc in diameter) were visible on the disk of the Sun.

2. *Measurement and results*

2.1. *Characteristic curves.*—The wedge spectrum and the "blank" spectrum of the centre of the Sun were run through the microphotometer (**14**) perpendicular to the direction of the dispersion. At D_3 the tracing was made at its wave-length, but at Hα two tracings were made, one on either side of and well clear of the absorption line. R. J. Bray's calibration of the step wedge (**1**) refers only to the wave-length range 5461–6143 A, but by his courtesy I was able to measure his plates also at the wave-length 6507 A. Within the errors of measurement the transmissions of the steps at this wave-length were identical with those found by Bray, so that the wedge may safely be regarded as strictly neutral in the wave-length range from D_3 to Hα. From these transmissions and the microphotometer tracings obtained as above, characteristic curves, corrected in the usual way for non-uniformity of illumination by means of the "blank" spectrum, were plotted giving microphotometer deflection as a function of the logarithm of the incident flux for a single plate. It was found that the curve for D_3 and the curves for Hα could be slid into coincidence so that a single well-determined characteristic curve for each plate could be used for both wave-lengths.

2.2. *Total intensities of chromospheric lines.*—The spectra of the chromosphere were marked in a dividing engine with scratches to the violet of D_3 and the red

of Hα. These scratches were so placed, allowing for the difference in magnification of the spectroscope at D_3 and Hα, that they corresponded to intervals of height of 1120 km for Series I (900 km for Series II and III) measured from a common though arbitrary origin for both lines. A chromospheric spectrum was run through the microphotometer between each pair of scratches corresponding to a common height. A typical low magnification (×6·9) tracing of a Series I chromospheric spectrum is shown in Fig. 1; the microphotometer slit-width is 0·011 mm and its height corresponded to approximately 1000 km in the chromosphere. In this tracing the peaks corresponding to the chromospheric emission

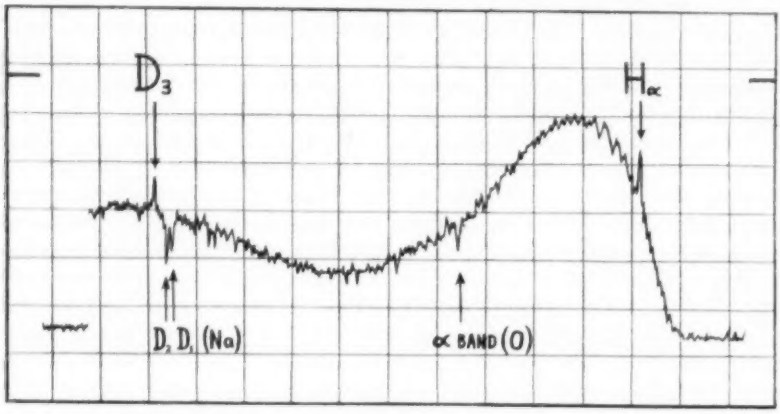


FIG. 1.—*Microphotometer trace of a Series I chromospheric spectrum, showing the two emission lines and the continuous spectrum of scattered light with absorption lines. Zero and clear film are at each end of the trace.*

lines D_3 and Hα and the continuous background due to solar light scattered in our atmosphere and by the telescope mirrors, are clearly seen. The microphotometer deflection at a point x on the plate at an emission line corresponds to a flux $K\{I(x)+S(x)\}$ where $I(x)$ is the flux at the point x in the emission line, $S(x)$ is the scattered light at the same position, and K is a constant required to convert the arbitrary fluxes of the characteristic curve into absolute units. From the microphotometer tracing of any one of the photospheric spectra we can similarly find $KJ(x)$ where $J(x)$ is the flux incident on the plate in that spectrum corresponding to the same point x. The ratio of these two quantities eliminates the common factor K. To find $S(x)/J(x)$ at D_3 we only need to measure the microphotometer deflection of the scattered continuous spectrum at the line. This was done by measuring the deflection corresponding to a line drawn on the trace at the level of the continuous spectrum across the emission line. At Hα, on the other hand, there is a well marked absorption line in the spectrum of scattered light, and measurement of the continuous spectrum at the line is not sufficient. The scattered light is predominantly due to photospheric limb light, as shown by A. B. Hart (**4**), and therefore from the limb spectrum on each of our plates together with the intensity of the continuous spectrum of scattered light we can find the profile of the absorption line in the scattered light, that is the true value of $S(x)/J(x)$ for Hα.

In this way we obtain for each line the quantity

$$i(x) = I(x)/J(x) = \{I(x) + S(x)\}/J(x) - S(x)/J(x) \tag{1}$$

where $i(x)$ is a number. The flux $J(x)$ received on the plate per unit interval at the point x in any of the photospheric spectra is given by

$$J(x) = \alpha(\lambda) J(\lambda) d\lambda/dx \tag{2}$$

where $J(\lambda)$ is the surface brightness of the centre of the Sun's disk in pure continuous spectrum in erg cm^{-2} sec^{-1} per A per unit solid angle. According to Mulders (**11**), $J(\lambda) = 3{\cdot}41 \times 10^6$ and $2{\cdot}88 \times 10^6$ at D_3 and Hα respectively. The factor $d\lambda/dx$ is the dispersion found from the Hartmann constants for the plate, and $\alpha(\lambda)$ is a factor giving the transmission ratio of any filters placed before the slit for the photospheric spectrum relative to those placed before it for the chromospheric spectrum. It also includes the darkening to the limb, taken from the results of Pierce, McMath, Goldberg, and Mohler (**13**), for the limb spectrum if it is used. We thus have for the total intensity of the chromospheric line expressed in erg cm^{-2} sec^{-1} per unit solid angle

$$\int I(x)dx = \alpha(\lambda) J(\lambda) d\lambda/dx \int i(x)dx \tag{3}$$

where the integral is taken over the emission line. The factors outside the integral sign are effectively constant over the width of the line.

The profile, $i(x)$, of a chromospheric line as observed with a spectroscope of small instrumental and photographic resolving power is wholly determined by the apparatus function. Hence the profile of such a line is always the same and to measure

$$\int i(x)dx$$

it is only necessary to measure the central intensity, $i(x_0)$, after establishing an empirical relation between the two quantities. For this purpose a number of

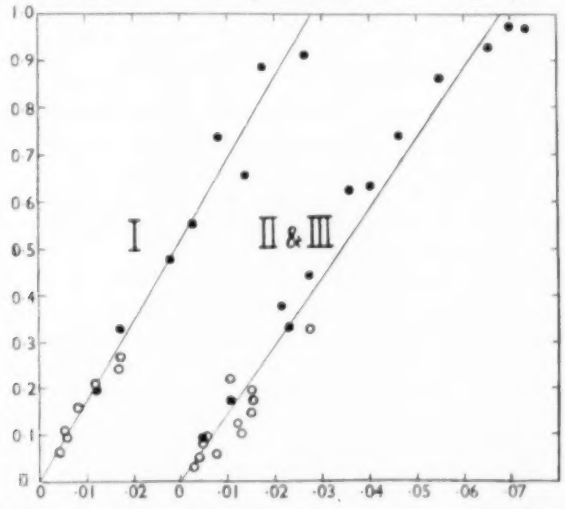


FIG. 2.—*Relation between central intensity, $i(x_0)$, and total intensity, $\int i(x)dx$. Points from D_3 profiles are marked with open circles, and those from Hα profiles with filled circles. Ordinates are central intensity and abscissae total intensity in mm.*

high magnification (×48·2) microphotometer tracings of chromospheric spectra were measured from which was found the relation, shown in Fig. 2,

$$\int i(x)dx = Ai(x_0) \tag{4}$$

where for Series I, $A = 0{\cdot}05810$ mm and, for Series II and III, $A = 0{\cdot}06794$ mm. The difference in the values of A arises from the fact that all Series I plates were traced with a microphotometer slit-width of 0·011 mm and all Series II and III plates with a slit-width of 0·057 mm. With these relations it now suffices to measure $i(x_0)$ on all microphotometer tracings in order to find total intensities of the chromospheric lines from equation (3).

2.3. *Position of the limb of the Sun.*—We thus obtain total line intensities for a number of chromospheric heights with an arbitrary origin for the height scale. To determine the true origin, which is the limb of the Sun, we must find its position in each chromospheric spectrum. For perfect seeing and instruments free from aberration the limb would appear on the plate as a sharp line. In fact there is a continuous fall of intensity from the spectrum of the photosphere across that of the chromosphere. If, following Wanders (**21**), we assume that a point object gives rise in the focal plane of the spectroscope to a flux distribution

$$(1/2\pi a^2) \exp(-s^2/2a^2),$$

where the parameter a may be taken to correspond to the radius of the tremor disk, we may readily calculate the continuous change of intensity across the limb. Since we are interested in points near the limb only, we can assume with sufficient accuracy that the surface brightness of the Sun near the limb may be represented by the linear relations

$$\begin{aligned} J(y) &= J_0(1 - my) & y &< 0 \\ &= 0 & y &> 0. \end{aligned}$$

Then Wanders shows that the brightness distribution across the limb is given by

$$j(y) = (1/\sqrt{\pi}).J_0\{(ma/\sqrt{2}) \exp(-y^2/2a^2) + (1 - my) \int_{y/a\sqrt{2}}^{-\infty} \exp(-u^2)du\}. \tag{5}$$

When measuring the central intensity, $i(x_0)$, of the chromospheric lines we measure $S(x_0)/J(x_0)$ for the continuous spectrum at several heights in the chromosphere. Because $j(y)$ is identical with the value of $S(x)$ at a height y above the limb of the Sun, we have from the values of $S(x_0)/J(x_0)$ the run of $j(y)$. In this way the observed $j(y)$ has the same arbitrary origin for its height scale as the chromospheric measures.

From equation (5) for a number of arbitrary values of the parameter, a, a set of curves was drawn, and of these for any spectrum that one was chosen which best fitted the observed curve. In this way both the position of the limb and the value of a were found for each chromospheric spectrum.

The above method of determining $j(y)$ was used only for Series I where, as already mentioned in Section 1, the limb spectrum was over-exposed. Thus the fit of theoretical and observed curves had to be made at the tail end of the limb curve where scattered light tends to falsify the result. There is therefore some possibility in this series that the limb has systematically been put too low by between 200 and 400 km. For Series II and III the parts of the limb near the chromosphere although taken through the steps of the wedge were in most cases

still slightly over-exposed, except to the red of Hα where plate sensitivity falls off rapidly. It was possible to make a measurable microphotometer trace perpendicular to the dispersion across the limb here, and such traces were used to determine $j(y)$. Connection in height scale with the chromospheric measures was made by using one of two shadows cast by wires on the step wedge.

TABLE I

Total intensities of chromospheric lines in units of 10^5 erg cm^{-2} sec^{-1} per unit solid angle

B	2700 km		4500 km	
	D_3	Hα	D_3	Hα
Series I ($a = 2''.91 \pm 0''.42$)				
−82	0.56	6.24	0.37	4.95
−81	0.66	6.03	0.40	4.94
−80	0.51	5.70	0.64	5.26
−78	0.71	6.27	0.56	4.84
−75	0.78	6.77	0.56	5.33
−72	0.68	6.92	...	4.94
−68	0.66	5.27	0.41	3.24
−67	...	4.14	...	3.85
−65	0.55	5.05	0.27	3.75
−59	0.73	5.55	0.56	4.07
−33	1.48	5.04	1.03	3.37
−15	0.94	4.97	0.64	3.91
− 3	1.19	5.26	0.57	2.79
5	0.85	4.90	0.53	2.81
Series II ($a = 2''.76 \pm 0''.79$)				
−78	0.48	5.44	0.22	2.84
−59	0.40	4.84	0.27	2.85
−51	0.60	4.75	0.25	2.09
−45	0.44	3.11	0.19	1.50
−37	0.60	4.65	0.29	2.53
−30	0.59	5.77	0.24	3.27
−23	0.62	4.35	0.30	2.12
15	0.68	4.21	0.39	2.82
22	0.68	4.31	0.39	2.55
28	0.94	5.62	0.46	2.45
37	0.86	4.52	0.51	2.76
50	0.74	3.97	0.42	2.28
Series III ($a = 3''.23 \pm 0''.67$)				
− 3	0.45	3.10	0.34	2.10
0	0.68	3.87	0.56	1.90
+ 9	0.69	3.82	0.44	2.10
+14	0.70	4.65	0.35	2.54
22	1.16	4.33	0.78	2.35
29	0.79	4.32	0.39	2.30
37	0.67	3.97	0.58	2.45
44	0.75	3.87	0.35	2.28
53	0.55	4.01	0.59	2.65
55	0.62	4.48	0.25	2.53
62	0.77	4.33	0.38	2.29
65	0.77	5.12	0.36	3.34

2.4. *Results.*—As we now know the total intensity of a chromospheric line for a number of heights above the limb, we can read from a curve for each chromospheric spectrum the total intensity at any required height. Results are shown in Table I for the three series of plates. The first column of this table gives the mean solar latitude of the three (Series I) or two (Series II and III) chromospheric spectra on each plate. Since these spectra were taken in fairly rapid succession the difference in latitude between successive spectra on a plate never amounts to more than one degree. The second and third columns give the total intensity of D_3 and Hα at a height of 2700 km above the limb, and the fourth and fifth the total intensity of the same two lines at a height of 4500 km. These total intensities are the means from the two or three chromospheric spectra on each plate and are expressed in units of 10^5 erg cm^{-2} sec^{-1} per unit solid angle.

3. *Latitude variation.*—First we consider Series I alone. The mean total intensities in Table I for this series are plotted in Fig. 3 as a function of heliographic latitude. The vertical lines through the points are the r.m.s. errors of a single observation as found from the two or three individually observed intensities at the latitude in question. As can be seen from the figure there is a great deal of scatter in the points which, as the r.m.s. errors indicate, is not entirely due to photometric error.

Assuming that the total intensity, I, of a line can be represented by the linear expression

$$I = p + q|B|$$

where B is the heliographic latitude, the values of p and q with their r.m.s. errors can be found by least squares. The individual values, not the means in Table I, are used for this least squares solution. The results are shown by straight lines in the figure, while the values of p and q are given in Table II. These lines slope up to the equator for D_3 showing stronger emission with lower latitudes, but for Hα slope in the opposite way indicating stronger emission towards the pole. For D_3 the slope, q, at 2700 km exceeds its r.m.s. error by a factor greater than three, and is therefore different from zero at better than the one per cent level of significance. At 4500 km, however, q is not significantly different from zero. For Hα the slope, q, is significantly different from zero at better than the one per cent level for both heights. Thus, on this day, superposed on a large random scatter not entirely due to observational error there was a tendency for Hα to be strengthened and D_3 weakened towards the pole, the Hα changes being more significant in that they were found at two heights above the limb.

Taking all three series together, we have, in Fig. 4, the intensities of Table I plotted as a function of heliographic latitude. The straight lines are the result of least squares solutions, using the mean values of Table I and treating each hemisphere separately. The values of p and q from these solutions are given in Table II. The slopes of the lines show the same variation for D_3 and Hα as was found for Series I alone, but only those for Hα in the southern hemisphere are significantly different from zero (one per cent level). Thus, for the average of the three days of sunspot minimum when the observations were made, Hα was significantly stronger towards the south pole compared with the equator, while in the northern hemisphere, within the scatter of the observations, was independent of latitude.

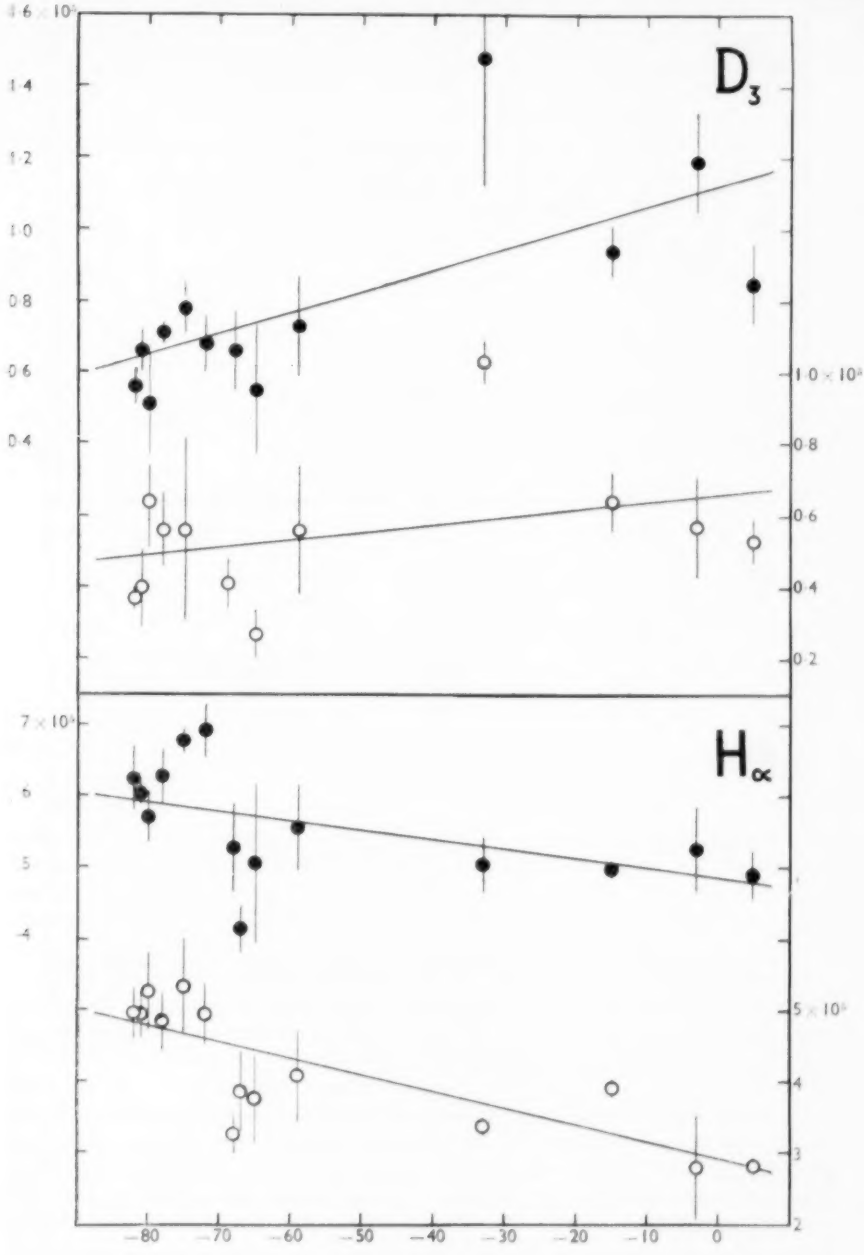


Fig. 3.—Variation of emission in D_3 and Hα with latitude, Series I. Ordinates are total energy in the line in erg cm^{-2} sec^{-1} per unit solid angle, and abscissae latitude in degrees. Measures at 2700 km and 4500 km above the limb are marked with filled and open circles respectively, and are displaced relative to each other in ordinates.

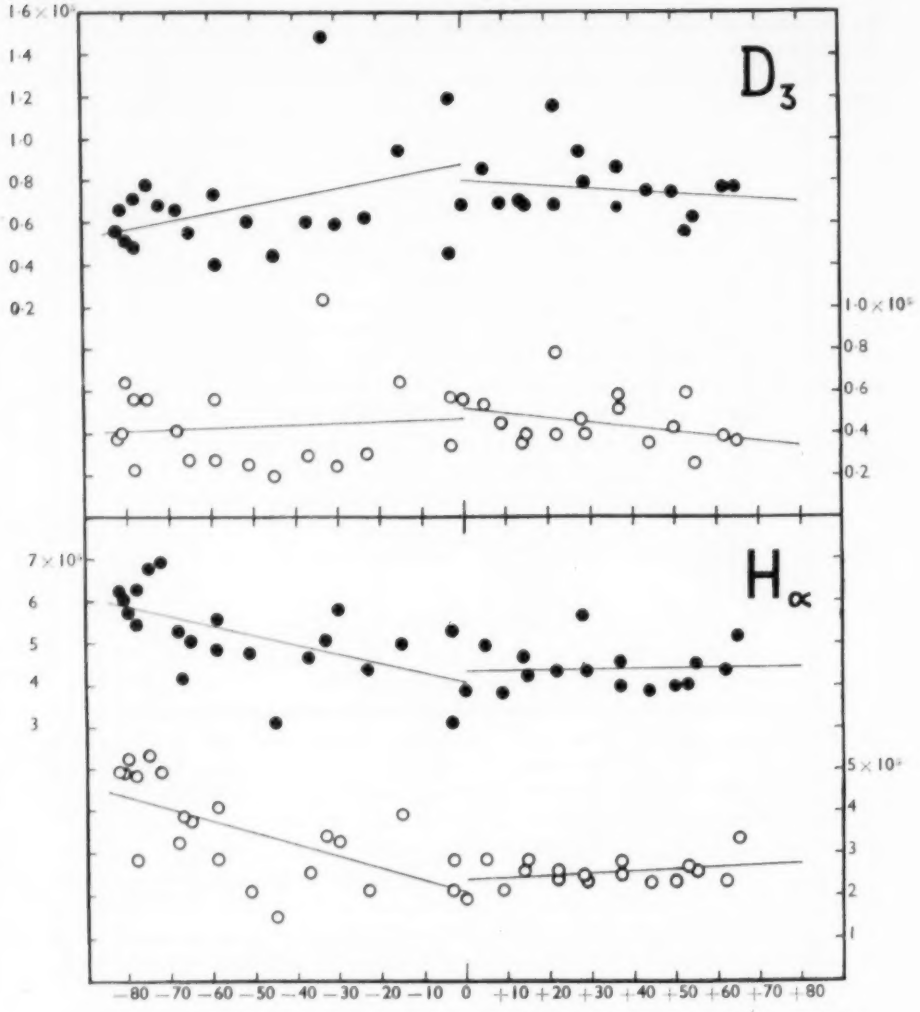


FIG. 4.—*Variation of emission in D_3 and Hα with latitude, Series I to III combined. Ordinates are total energy in the line in erg cm^{-2} sec^{-1} per unit solid angle, and abscissae latitude in degrees. Measures at 2700 km and 4500 km above the limb are marked with filled and open circles respectively and are displaced relative to each other in ordinates.*

The different linear latitude dependence found for Series I on the one hand, and for the three series together on the other, suggests that the dependence on latitude may vary. To look for day to day changes we consider the series separately and in the least squares solutions use individual observations. In the southern hemisphere consider Series I and Series II. For D_3 we find a slope up to the equator for those points of Series II in this hemisphere as was found for Series I. The value of q at 2700 km is significantly different from zero at the ten per cent level. Thus we have the same variation as Series I, but the general level of emission for Series II is lower and combining the two, as in taking

all three series together, increases the scatter and the significance is lost. For Hα in this hemisphere we find that Series II gives a slope which is not significantly different from zero and differs from that given by Series I by more than the sum of their r.m.s. errors. Evidently there is a change in the linear variation between the two days. Series III taken alone shows an increase in Hα emission to the north pole, with q significantly different from zero at the one per cent level at

TABLE II

Parameters for the least-squares solutions

Spectra	Height (km)	D_3		Hα	
		p	q	p	q
Series I–III (−ve lat.)	2700	$0{\cdot}87 \pm 0{\cdot}06$	$-0{\cdot}0036 \pm 0{\cdot}0022$	$4{\cdot}04 \pm 0{\cdot}19$	$0{\cdot}0221 \pm 0{\cdot}0077$
	4500	$0{\cdot}46 \pm 0{\cdot}05$	$-0{\cdot}0007 \pm 0{\cdot}0018$	$2{\cdot}05 \pm 0{\cdot}20$	$0{\cdot}0286 \pm 0{\cdot}0083$
Series I–III (+ve lat.)	2700	$0{\cdot}80 \pm 0{\cdot}03$	$-0{\cdot}0012 \pm 0{\cdot}0018$	$4{\cdot}32 \pm 0{\cdot}12$	$0{\cdot}0016 \pm 0{\cdot}0060$
	4500	$0{\cdot}52 \pm 0{\cdot}03$	$-0{\cdot}0020 \pm 0{\cdot}0015$	$2{\cdot}33 \pm 0{\cdot}08$	$0{\cdot}0052 \pm 0{\cdot}0038$
Series I	2700	$1{\cdot}12 \pm 0{\cdot}04$	$-0{\cdot}0059 \pm 0{\cdot}0013$	$4{\cdot}83 \pm 0{\cdot}12$	$0{\cdot}0137 \pm 0{\cdot}0041$
	4500	$0{\cdot}66 \pm 0{\cdot}10$	$-0{\cdot}0021 \pm 0{\cdot}0021$	$2{\cdot}89 \pm 0{\cdot}10$	$0{\cdot}0236 \pm 0{\cdot}0034$
Series II (−ve lat.)	2700	$0{\cdot}69 \pm 0{\cdot}03$	$-0{\cdot}0036 \pm 0{\cdot}0018$	$4{\cdot}59 \pm 0{\cdot}09$	$0{\cdot}0017 \pm 0{\cdot}0059$
	4500	$0{\cdot}29 \pm 0{\cdot}01$	$-0{\cdot}0008 \pm 0{\cdot}0009$	$2{\cdot}41 \pm 0{\cdot}20$	$0{\cdot}0009 \pm 0{\cdot}0134$
Series III	2700	$0{\cdot}69 \pm 0{\cdot}04$	$0{\cdot}0009 \pm 0{\cdot}0018$	$3{\cdot}73 \pm 0{\cdot}11$	$0{\cdot}0131 \pm 0{\cdot}0046$
	4500	$0{\cdot}49 \pm 0{\cdot}04$	$-0{\cdot}0016 \pm 0{\cdot}0016$	$2{\cdot}07 \pm 0{\cdot}07$	$0{\cdot}0103 \pm 0{\cdot}0031$

both heights, but shows no significant change for D_3. Thus in the southern hemisphere we find a variation between Series I and II in the level of D_3 emission and in the slope of the linear change of Hα emission, and in the northern hemisphere an increase in Hα emission towards the pole on one day, the significance of which is lost by combining with a few measures on other days.

Thus if we assume that the intensities show an underlying *linear* dependence on latitude, the observations of all three series can be reconciled only if we further assume that this dependence is a function of time and of hemisphere. However, in Series I, which shows the most clearly marked linear dependence, the observations for $|B| < 70°$ are few, irregularly distributed in latitude and show a large scatter. These observations could thus be interpreted equally well on the assumption that in the latitude range $0 \leqslant |B| < 70°$ the emission, apart from local irregularities, is constant, and only for $|B| > 70°$ is there a *polar* strengthening for Hα and a *polar* weakening for D_3. Since for Series II there is only one observation for which $|B| > 70°$ and for Series III none, it is not surprising on the hypothesis of a purely polar effect that these two series show no pronounced evidence of a latitude effect. Thus the results from all three series are reconciled on the hypothesis of a purely polar effect, and there is no necessity for assuming that this effect is variable with time or with hemisphere.

The spectrophotometric measures of intensity obtained in this investigation are clearly not sufficiently numerous to enable us to distinguish between the hypothesis of a linear dependence of emission on latitude and the hypothesis of a purely polar effect. Unfortunately the non-photometric estimates of chromospheric height, made at Arcetri (**3**) and Tokyo (**16**), though of interest in showing similar latitude effects, are of no help in settling this point because of their uncertain physical interpretation.

Mention has been made of local irregularities in the chromospheric emission. A particular interesting example occurs at latitude $B = -33°$ in Series I. Though the r.m.s. error of a single measure is large there seems little question that D_3 was abnormally strong at this latitude at this time, but that the simultaneously observed Hz showed a normal intensity. So strong is the emission in D_3 that we may be observing, in the phraseology of Cillié and Menzel (**2**), a "hot spot" where the excitation temperature of helium is abnormally high. If this be so it is remarkable that the hydrogen emission is unaffected. On the other hand the hydrogen emission may be affected, but the hot spot may occur at too large an optical depth in Hz to be visible in that wave-length.

4. *Height variation.*—To correct the observed intensity-height distributions for the effects of seeing we must solve the integral equation

$$O(y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} T(\xi)U(\eta, \xi - y)d\eta d\xi \qquad (6)$$

for $T(y)$, the true distribution, where $O(y)$ is the observed and $U(\eta, \xi)$ is the flux distribution in the focal plane of the spectroscope due to a point object as assumed in Section 2.3 with $s^2 = \xi^2 + \eta^2$. ξ and η are measured at right angles and parallel to the limb respectively.

Van de Hulst (**5**) has pointed out that in this type of equation random errors in $O(y)$ lead to irregularities in the determined $T(y)$ which are not really there. Kreisel (**10**) has shown that a rapidly oscillating $T(y)$ would have a small $O(y)$ and, since the equation is linear, solutions are superposable. The errors in $O(y)$ imply therefore that there are many functions $T(y)$ which will give the same $O(y)$ within the experimental error.

The experimental error in the observed $O(y)$ from a single chromospheric spectrum was considered to be too large for the determined $T(y)$ to be of any value, therefore the mean $O(y)$ over several spectra were used. This mean was found by reading from the total intensity against height curve for each spectrum the intensities at intervals of 900 km in height with the limb as origin, and taking the mean at each height. The method of solution of equation (6) was to assume a $T(y)$, calculate the corresponding $O(y)$, and seek agreement with the observations. This gave results lying more within the experimental error than those by other methods of solution. The type of curve assumed for $T(y)$ was of the form

$$\begin{aligned} T(y) &= E_0 \exp(-\beta y) && y > 0 \\ &= E_0 \exp(-\beta y) \int_{\sqrt{-\beta y}}^{+\infty} \exp(-u^2)du && y < 0. \end{aligned} \qquad (7)$$

This is the height distribution for the emission in the line of sight from a chromosphere with emission per unit volume following a height variation law

$$I(h) = I_0 \exp(-\beta h)$$

and without self-absorption. Substituting in equation (6) gives a theoretical $O(y)$ which can be fitted to the observations by varying E_0 and β. This was done to obtain a least sum of squares of the intensity differences between calculated $O(y)$ and the observed intensities.

The results are shown in Table III and are confined to Series II and III which, as already explained, give a better height range than Series I. The first column gives the spectra used to obtain the mean $O(y)$, the second and third give the values of E_0 and β found for the true distribution, $T(y)$, of the form of equation (7). E_0 is in units of 10^5 erg cm^{-2} sec^{-1} per unit solid angle and β in units of 10^{-8} cm^{-1}. The fourth column gives the mean radius of the tremor disk in seconds of arc with its r.m.s. error. The Hα height variations of Series III were divided into four groups according to latitude to seek any latitude variation effects. The last group, those nearest the pole, shows a lower value of β but not too much significance can be attached to this without further investigation. The Series II Hα results were not subdivided according to latitude as the local variations are greater than any overall latitude changes, and the greater number of spectra in the mean gives a better defined $O(y)$. For the same reason no attempt has been made to subdivide the D_3 results according to latitude.

In Fig. 5 we have the points of the observed $O(y)$ of the last four, i.e. complete series, groups of the first column of Table III together with a curve showing the computed $O(y)$ which is the least square solution. This is the $O(y)$ corresponding to the values of E_0, β, and a of the table.

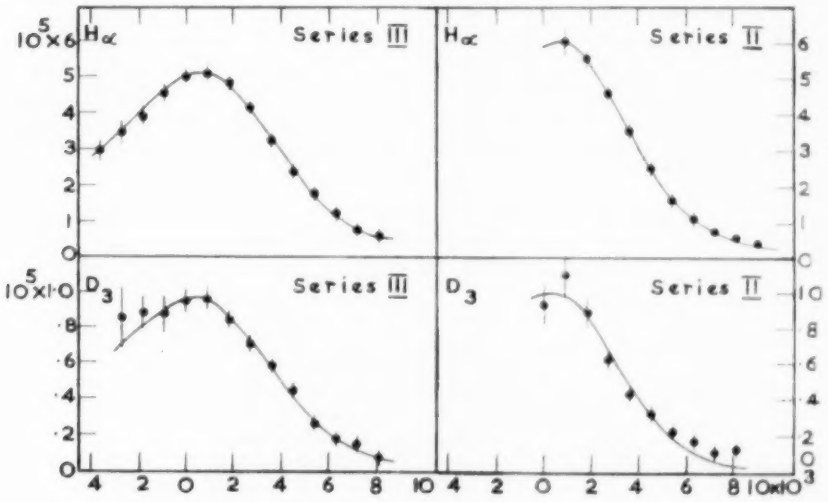


FIG. 5.—*Observed height distribution of D_3 and Hα emission. Vertical lines through the points are r.m.s. errors. Ordinates are total energy in the line in erg cm^{-2} sec^{-1} per unit solid angle, and abscissae height above the limb of the Sun in km.*

The figure shows that, with the exception of D_3 Series II, the observations are in good agreement with an exponential height distribution of emission in the chromosphere, and the different groups of spectra for Hα give consistent values for β. With such large values for the radius of the tremor disk we cannot hope to get more than an average true height distribution for the emission. Local

TABLE III

Spectra	E_0	β	a
Hα Series III lat. 0° to 9°	12·2	0·49	3·39 ± 0·32
lat. 14° to 29°	15·7	0·53	3·47 ± 0·22
lat. 37° to 53°	14·2	0·50	3·32 ± 0·33
lat. 55° to 65°	13·3	0·42	2·75 ± 0·09
Series III (all)	13·7	0·48	3·23 ± 0·67
Series II	15·1	0·49	2·76 ± 0·79
D_3 Series III	2·87	0·58	3·23 ± 0·67
Series II	2·95	0·67	2·76 ± 0·79

irregularities must exist in the nature of spicules, but will be missed by the above methods.

The height variation of D_3 and Hα emission have been measured at the 1932 eclipse by Cillié and Menzel (**2**) and at the 1941 and 1945 eclipses by Vyazanitsyn (**19**) (**20**). They measured the total emission in the lines from a 1 cm wide strip of the chromosphere above a given height and found an exponential distribution for this. Differentiating gives an exponential distribution, with the same gradient, β, for the variation of the emission per cm^{-2}, which is what we have measured. In Table IV are the gradients β, of these three eclipse observations for D_3 and Hα, given in units of 10^{-8} cm^{-1}.

Outside of an eclipse, exponential gradients for D_3 and Hα emission have been found by V. A. and T. V. Krat (**8**) (**9**). The observations were made during 1951, 2 and 3. We have placed these values of β in Table IV together with the eclipse observations. Those for D_3 are Krat's values for the chromosphere above inactive regions of the sun. The values above active regions are lower.

TABLE IV

Exponential Gradient, β, for Emission

		D_3	Hα
Eclipse	1932	1·33	1·16
	1941	0·88	0·90
	1945	0·73	0·86
Non-eclipse	1951	0·71	...
	1952	0·55	0·53
	1953	...	0·52

Our results are consistent with the non-eclipse observations, but the eclipse values all appear greater. A possible explanation of this is that the chromosphere has actually changed, or it may be that the height distribution is only approximated by an exponential function and the two entirely different methods give different results when an exponential distribution is assumed. Differences are seen between the eclipse observations so both possibilities might each have an effect.

Conclusion.—This investigation has necessarily been of an exploratory character. Nevertheless, underlying a large random scatter not wholly due to photometric error, there is clear evidence of a dependence of chromospheric emission on solar latitude. This is such that Hα is strengthened and D_3 weakened at the poles. The observations, however, are not in themselves sufficient to distinguish between on the one hand a linear dependence on latitude, possibly varying from day to day and between hemispheres, and on the other a purely polar strengthening (or weakening) with no dependence on latitude in the non-polar regions. To distinguish between these alternatives will require a very extended series of observations simply because of the large local variations in chromospheric emission.

Since the principal purpose of the investigation was to find if there were any latitude dependence, observations were made on days with long clear periods, rather than at those rare intervals characterized by good seeing. The observations are therefore not well suited for a determination of the dependence of emission on height. Nevertheless the discussion in Section 4 has not been wholly unsuccessful, and there appear to be good grounds for expecting that non-eclipse observations made at times of good seeing may yield an accurate distribution of emission with height.

I wish to express my gratitude to Professor H. H. Plaskett for his unfailing encouragement and advice and his generous assistance, particularly in making the observations, and to thank the Department of Scientific and Industrial Research for a maintenance grant, during the tenure of which the work has been carried out.

University Observatory,
Oxford:
1956 *September* 17.

References

(1) R. J. Bray, *M.N.*, **114**, 547, 1954; Comm. Univ. Obs., Oxford, No. 45.
(2) G. G. Cillié and D. H. Menzel, Harvard Circ. 410, 1935.
(3) M. G. Fracastoro, *Oss. di Arcetri*, **64**, 44, 1948.
(4) A. B. Hart, *M.N.*, **114**, 275, 1954; Comm. Univ. Obs., Oxford, No. 43.
(5) H. C. Van de Hulst, *B.A.N.*, **10**, 75, 1946.
(6) C. de Jager, *Rech. Astr. Obs. Utrecht*, **13**, Part I.
(7) P. C. Keenan, *Ap. J.*, **75**, 277, 1932.
(8) T. V. Krat, *Izv. Glavn. Astr. Obs. Pulkovo*, No. 152, **19**, 20, 1954.
(9) V. A. Krat and T. V. Krat, *Izv. Glavn. Astr. Obs. Pulkovo*, No. 153, **19**, 1, 1955.
(10) G. Kreisel, *Proc. Roy. Soc.* A, **197**, 160, 1949.
(11) G. W. F. Mulders, *Zs. f. Ap.*, **11**, 132, 1935.
(12) E. J. Perepelkin and O. A. Melnikov, *Bull. Obs. Central Pulkovo*, No. 122, **14**, 1, 1935.
(13) A. K. Pierce, R. R. McMath, L. Goldberg and O. C. Mohler, *Ap. J.*, **112**, 289, 1950.
(14) H. H. Plaskett, *M.N.*, **95**, 160, 1934; Comm. Univ. Obs., Oxford, No. 4.
(15) H. H. Plaskett, *M.N.*, **115**, 542, 1955; Comm. Univ. Obs., Oxford, No. 52.
(16) K. Sotome, Tokyo Astr. Bull., No. 90, 1933.
(17) A. Unsöld, *Zs. f. Phys.*, **59**, 372, 1929.
(18) A. Unsöld, *Zs. f. Ap.*, **3**, 77, 1931.
(19) V. P. Vyazanitsyn, *Izv. Glavn. Astr. Obs. Pulkovo*, No. 147, **18**, 19, 1951.
(20) V. P. Vyazanitsyn, *Izv. Glavn. Astr. Obs. Pulkovo*, No. 149, **19**, 40, 1952.
(21) A. J. M. Wanders, *Zs. f. Ap.*, **8**, 111, 1934.

A MODEL ATMOSPHERE FOR THE SOLAR LIMB BASED ON CONTINUUM OBSERVATIONS

B. E. J. Pagel

(Communicated by the Astronomer Royal)

(Received 1956 December 8)

Summary

The photospheric temperature distribution previously derived by the author is combined with the results for the low chromosphere obtained by the High Altitude Observatory group from observations of the continuum at the 1952 eclipse. The resulting model, assumed to be in hydrostatic equilibrium with negligible turbulence and a hydrogen–helium ratio 10 : 1, is used to calculate the intensities of continuous radiation at various heights near the limb and in the lowest 600 km of the chromosphere at $\lambda\lambda$ 4700, 3650, and 3640, carrying out the numerical integrations with the aid of EDSAC. The gradient and absolute values of the emission at λ 4700 are in satisfactory agreement with observations in the lowest part of the chromosphere, as is the general behaviour of the calculated emission in the Balmer continuum.

The darkening at the extreme limb is calculated, taking curvature into account, from the adopted model and from the non-grey radiative-equilibrium model calculated by Böhm (1954 a). The observational results obtained at eclipses scatter too much among themselves for a convincing comparison to be carried out.

The centre-limb variation in the Balmer discontinuity agrees well with observations for $\cos\theta \leqslant 0{\cdot}5$ and satisfactorily for $\cos\theta > 0{\cdot}5$. This suggests equilibrium excitation of hydrogen to the second quantum level in those regions where Balmer continuum absorption is significant ($\tau_0 > 0{\cdot}001$).

---

1. *Introduction.*—For the interpretation of Fraunhofer and flash spectra, and also in order to obtain some insight into the origin of the chromosphere, it is necessary to know the variation with height of physical conditions near the solar limb. The core of the problem consists in finding the kinetic temperature as a function of optical depth; for, when this is known, it is possible to deduce the pressure and electron pressure from the principle of hydrostatic equilibrium, the applicability of which to the limb region is established below (Section 3). We assume for this purpose (*a*) that the electrons at a given level in the solar atmosphere have a Maxwellian distribution of velocities characterized by a definite kinetic temperature T, and (*b*) that T may be identified with the gas-kinetic temperature. The justification for (*a*) has been given by Woolley and Stibbs (1953); that for (*b*) is due to Bhatnagar, Krook, Menzel and Thomas (1955).

The temperature distribution in a model atmosphere is generally either deduced from a theoretical principle (e.g. radiative equilibrium) or inferred from observational data in the continuum or in Fraunhofer lines. The theoretical method is unsuited to our present purpose because the presence or absence of a non-radiative energy source in the low chromosphere constitutes one of the prime questions which the determination of an empirically correct temperature distribution should help to solve. For this reason, we shall base ourselves

exclusively on observational data. Furthermore, since the mechanism of re-emission in Fraunhofer lines is in general a highly complex one, we shall make exclusive use of observational data in the continuum to obtain a model solar atmosphere for $\tau_0 \geqslant 10^{-5}$, where τ_0 is the monochromatic optical depth at $\lambda 5000$; for $\tau_0 < 10^{-5}$, the assumption of hydrostatic equilibrium is likely to break down (Section 3). The model is verified by calculating the continuous emission as a function of position on and off the solar disk near the limb, the effects of curvature and self-absorption being accurately allowed for by high-speed automatic computation, and comparing the results with measurements carried out at solar eclipses.

Making assumption (*a*) above, but without further restrictions if the densities of the emitting particles are known, the monochromatic continuous emission from unit volume is given by

$$j_\lambda = \rho k_\lambda^{(T)} B_\lambda(T), \tag{1}$$

where ρ is the mass density, $k_\lambda^{(T)}$ the mass absorption coefficient computed assuming thermodynamic equilibrium at the electron temperature T, and $B_\lambda(T)$ is the corresponding Planck emission function (Zanstra, 1950; Pagel, 1957). $k_\lambda^{(T)}$ may be readily calculated by combining the effective absorption cross-section due to H^- given by Chandrasekhar and Breen (1946) with the standard Kramers–Gaunt formula for the cross-section of neutral hydrogen; but the true absorption coefficient k_λ is liable to be affected in the outer solar layers by departures from thermodynamic equilibrium, particularly in the ionization of H^-. If x is the degree of dissociation of H^- in the presence of N_e electrons per unit volume, we have

$$\left(\frac{xN_e}{1-x}\right) \Big/ \left(\frac{xN_e}{1-x}\right)^{(T)} = \frac{N_e S + Q}{N_e S + Q^{(T)}}, \tag{2}$$

where $N_e S$ is the probability of dissociation of an H^- ion in unit time by collision with electrons, Q is the probability of photoelectric dissociation, and the superscript $^{(T)}$ refers to local thermodynamic equilibrium at temperature T. The quantities Q, S are given by

$$Q = 4\pi \int_{\nu_0}^{\infty} J_\nu \psi_\nu \, d\nu = \frac{4\pi}{h} \int_{\nu_0}^{\infty} \alpha_{BF} \frac{J_\nu}{\nu} \, d\nu \tag{3}$$

where J_ν is the radiation intensity averaged over all directions, ν_0 is the frequency corresponding to the dissociation potential of H^- ($h\nu_0 = \epsilon_0 = 0{\cdot}747$ eV), ψ_ν is the probability that a given photon will be absorbed by a given H^- ion, and α_{BF} is the photo-ionization cross-section; and

$$S = \int_{v_0}^{\infty} v M(v) \sigma'(v) \, dv = \frac{1}{m} \int_{\epsilon_0}^{\infty} M(v) \sigma(\epsilon) \, d\epsilon \tag{4}$$

where $M(v)$ is the Maxwell distribution function of velocity, $\sigma'(v)$, $\sigma(\epsilon)$ are the inelastic collision cross-sections expressed as functions of velocity and energy respectively, and m is the mass of an electron.

The cross-section $\sigma(\epsilon)$ for H^- is not precisely known; an order-of-magnitude estimate may be obtained by assuming hydrogen-like behaviour and using the classical formula of Thomson (1912):

$$\sigma(\epsilon) = 3f \frac{\pi e^4}{\epsilon} \left(\frac{1}{\epsilon_0} - \frac{1}{\epsilon}\right) = \frac{3mce^2}{2h} \frac{\epsilon^2}{\epsilon_0^2} \alpha_{BF} \left(\frac{1}{\epsilon_0} - \frac{1}{\epsilon}\right)$$

since

$$\alpha_{BF} = f\frac{2\pi e^2}{mc}\frac{\epsilon_0{}^2 h}{\epsilon^3}$$

(Woolley, 1934). Equation (4) may then be written

$$S = 2{\cdot}44 \times 10^{10}\left(\frac{T}{4334}\right)^{3/2}\int_{2(4334/T)}^{\infty} u^3\left[\frac{1}{2}\left(\frac{T}{4334}\right) - \frac{1}{u}\right]\alpha_{BF}(u)e^{-u}\,du$$

where

$$u = \epsilon/kT.$$

α_{BF} was obtained from the calculations of Chandrasekhar (1945) by including a stimulated emission factor $(1 - e^{-u})$. Numerical integration then gave

$$S = 1{\cdot}1 \times 10^{-6}\,\text{cm}^3\,\text{sec}^{-1} \quad \text{for} \quad T = 4334$$

and

$$S = 2{\cdot}7 \times 10^{-6}\,\text{cm}^3\,\text{sec}^{-1} \quad \text{for} \quad T = 5800.$$

Q was calculated using the values of J_ν near the solar limb given by Böhm (1954 b), leading to

$$Q = 1{\cdot}26 \times 10^6\,\text{sec}^{-1}.$$

Q and $N_e S$ are thus of the same order of magnitude when

$$N_e = 1{\cdot}15 \times 10^{12} \quad \text{for} \quad T = 4334$$

and

$$N_e = 0{\cdot}47 \times 10^{12} \quad \text{for} \quad T = 5800.$$

An electron density of $10^{12}\,\text{cm}^{-3}$ corresponds to $\tau_0 \simeq 0{\cdot}01$. Below this layer, one may expect collisional dissociation to predominate, so that the Saha equation is approximately satisfied; at greater heights in the solar atmosphere, photo-ionization and collisional dissociation may be expected to be of the same order of magnitude owing to the low electron density. This can be shown by a reference to the values of T and of $\log P_e$ given in Table II, which are such that the rise of S with temperature, given above, just about compensates the decrease in N_e with increasing height, keeping $N_e S$ approximately constant and equal to Q. Since S is only roughly known, we do not calculate it in detail as a function of temperature, but write instead

$$\frac{N_{H^-}}{N_{H^-}{}^{(T)}} = \frac{1}{2}\left[\frac{Q^{(T)}}{Q} + 1\right] = \frac{1}{2}\left[\frac{\int\psi(\nu)B_\nu(T)\,d\nu}{\int\psi(\nu)J_\nu\,d\nu} + 1\right] \tag{5}$$

since $x \simeq 1$. The resulting correction factors to the effective bound-free absorption cross-sections due to H^- per hydrogen atom per unit electron pressure are given for four different temperatures in Table I. With the aid of Table I, the normal model-atmosphere method is applicable to chromospheric regions in hydrostatic equilibrium.

TABLE I

Departure from equilibrium in ionization of H^- in low chromosphere for $N_e \leqslant 10^{12}$

T	4100	4500	5000	5500
$\frac{N_{H^-}}{N_{H^-}{}^{(T)}}$ from equation (5)	0·80	0·90	1·08	1·37
$\log(Q/S)$	12·1	12·0	11·9	11·75

2. *The choice of a temperature distribution.*—The presence of granules in the photosphere and of spicules in the high chromosphere suggests that there must

be inhomogeneities in temperature and density at any given height in the atmosphere of the quiet Sun. However, there seems at present to be no evidence that the turbulence spectrum in the region between $\tau_0 = 0{\cdot}2$ and 1000 km above the limb can be approximated by a finite number of columns with different temperatures and velocities (Athay and Menzel, 1956; Voigt, 1956), and it therefore appears advisable to proceed as far as possible with an homogeneous model. Granting this, one is faced with the problem of choosing among the bewildering variety of photospheric temperature distributions put forward in recent years; a summary of these models is given by de Jager (1955), and, as de Jager points out, the differences between them may be attributed to departures from thermodynamic equilibrium in the outer layers of the Sun. This comment could perhaps be made more precise by pointing out that the "temperatures" found by the various authors concerned are actually different quantities. For instance, the low surface temperature ($<$4000 deg. C) inferred by Böhm-Vitense (1954, 1955) from the central intensities of the strongest Fraunhofer and flash lines is a "radiation temperature" formally definable by writing $B_\nu(T_r)$ for the source function in the transfer equation; physically, T_r represents the excitation temperature of the upper level of the transition in those regions (probably in the chromosphere) from which resonance radiation reaches the observer. The identification of T_r with any other sort of temperature, e.g. electron temperature, rests on the assumption that re-emission in the centres of resonance lines takes place according to "pure absorption" in local thermodynamic equilibrium; this assumption is, to say the least, open to question (see for example Plaskett, 1955). The same objection may be raised against taking too literally the low surface temperature calculated by Böhm (1954a) from blanketing-effect theory, though the effect in this case is not likely to be large; more serious is the fact that the theory of radiative equilibrium in a plane-parallel atmosphere is probably not applicable to the chromospheric regions that make the chief contribution to the cores of strong Fraunhofer lines. Reservations may also be expressed regarding the temperature distribution derived from the Balmer lines (de Jager, 1952), which are strongly affected by temperature inhomogeneities (de Jager, 1954) and by uncertainties in the absorption profile resulting from Stark broadening. The question of the excitation of hydrogen to the second quantum level is discussed below (Section 5). Thus, the only investigations bearing directly on electron temperatures in the smoothed-out approximation to the solar atmosphere are those relating to intensities in the continuum, though there are arguments for the view that the source function may approximate to $B_\nu(T)$ in faint and medium subordinate lines when these are excited mainly by radiation in frequencies for which the solar atmosphere is optically thick (Unsöld, 1955). A detailed study of a group of infra-red Fe I lines with excitation potentials in the neighbourhood of 5·60 eV led the writer to put forward a model in which the temperature actually reaches a minimum value at $\tau_0 = 0{\cdot}015$ and then rises towards the limb (Pagel, 1955); a similar behaviour of T was inferred at about the same time by Athay, Menzel, Pecker and Thomas (1955) from continuum measurements in flash spectra obtained at the 1952 eclipse. The general similarity between these two results as distinct from other estimates of temperature near the solar limb is a small piece of evidence in favour of the "pure absorption" approach to subordinate lines. It is also of interest that the minimum temperature

of 4100 deg. obtained by Pagel (1955) agrees well with the lowest excitation temperature of 4150 ± 50 deg. found by Menzel, Baker and Goldberg (1938) from Fe I and by van Regemorter (1953) from V I.

Fig. 1 shows the electron temperatures in the neighbourhood of the solar limb derived by Athay, Menzel, Pecker and Thomas (1955) and by Pagel (1955). It is clear that the agreement between the actual temperatures derived is poor; however, it will be shown that the temperature distribution illustrated by the continuous line is consistent with the available observational material. The physical interpretation of the model is beyond the scope of this paper; however, if its reality should be confirmed by further observations, the sudden inversion of the temperature gradient will have a strong bearing on theories of the origin of the chromosphere.

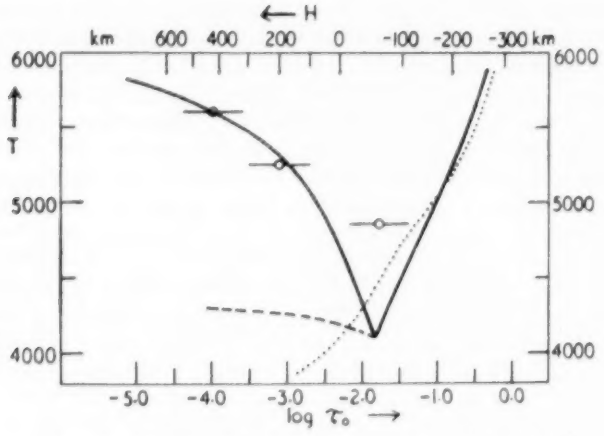


FIG. 1.—*Temperature distribution near the solar limb.*
Full line: adopted model.
Broken line: preliminary model (Pagel, 1955).
Dotted line: radiative-equilibrium model (Böhm, 1954 a).
Circles: temperatures derived by Athay et al. in height ranges indicated by horizontal lines.
H represents the geometrical height above the limb. This scale applies to the adopted model only.

3. *The calculation of a model atmosphere.*—Given the functional dependence of temperature on monochromatic optical depth, the total pressure P_g, electron pressure P_e, and height H may be deduced at each level of an atmosphere of given chemical composition by an iterative procedure if hydrostatic equilibrium is assumed (see, for example, Aller, 1953). In the neighbourhood of the solar limb, however, some authors have suggested that the material is supported against gravity by turbulent velocities of the order of 10 km/sec (McCrea, 1929; Unsöld, 1953; a lower value is suggested by de Jager, 1955) because the observed intensity gradient (Lindblad, quoted by Wildt, 1947; Lindblad and Kristenson, 1953) is considerably lower than would be expected in an isothermal atmosphere at 4000–5000 deg. This interpretation is open to grave doubt because such high velocities at this level would presumably manifest themselves in the results of line-width measurements. In fact, the vertical turbulent velocity parameter found from Fraunhofer lines is of the order of only 2 km/sec (Minnaert, 1953)

while the horizontal parameter at the limb has about the same value, though it rises to 16 km/sec at a height of 1600 km (Redman and Suemoto, 1954).* Since the corresponding thermal velocity parameter for hydrogen atoms at 4500 deg. is 8·7 km/sec, it follows that turbulent kinetic energy can make only a negligible contribution to the lift up to a few hundred km at least. We therefore proceed on the assumption of hydrostatic equilibrium.

In the calculation of the model atmosphere, a hydrogen–metal atomic abundance ratio $A = 10^4$ and a helium–hydrogen ratio $B = 0{\cdot}1$ were assumed. The fundamental parameters are τ_0, the optical depth at λ 5000, and the electron (or kinetic) temperature T. The bound–free opacity per unit electron pressure due to H^- given by Chandrasekhar and Breen (1946) was modified for $\tau_0 < 0{\cdot}006$ by the factors given in Table I. For $\tau_0 > 0{\cdot}006$, the bound–free opacity was assumed to have its equilibrium value, as were the free–free opacity of H^- and the opacity of neutral H at all depths. The validity of the latter assumption is discussed later (Section 5).

In spectral regions where H^- absorption predominates, the relationship between height and monochromatic optical depth is relatively insensitive to the assumed values of the cross-section $a(H^-)$ per hydrogen atom per unit electron pressure. For, in the region where metals contribute most of the free electrons

$$\frac{d\tau}{dH} = k\rho = \frac{a(H^-)P_e P_g}{(1+B)kT} \simeq \frac{a(H^-)P_g^2}{A(1+B)^2 kT}. \tag{6}$$

Following Aller (1953), we have

$$P_g^2(\tau) = 2gm_H A(1+B)(1+4B)\int_0^\tau \frac{dt}{a(H^-)}, \tag{7}$$

where m_H is the mass of a hydrogen atom. Hence, if $a(H^-)$ is changed by a constant factor, $d\tau/dH$ is unchanged. Similarly, $d\tau/dH$ is insensitive to the assumed value of A, but quite sensitive to that of B.

The computation of P_g and P_e was carried out by means of the iterative formula

$$P_g^{(1)}(\tau_0) = \left[2gm_H(1+4B)\int_0^{\tau_0} \frac{P_g^{(0)}\,dt_0}{a(H^-)P_e^{(0)} + a(H) + \frac{1}{2}\sigma P_e/P_g}\right]^{1/2} = 3{\cdot}58\times 10^{-10}\left[\int_0^{\tau_0} \frac{P_g^{(0)}\,dt_0}{a(H^-)P_e^{(0)} + a(H) + \frac{1}{2}\sigma P_e/P_g}\right]^{1/2}, \tag{8}$$

where $a(H)$ and $\frac{1}{2}\sigma P_e/P_g$ represent the contributions at λ 5000 of neutral hydrogen and Thomson scattering respectively. The factor $\frac{1}{2}$ is a geometrical dilution factor, the use of which is justified by the fact that electron scattering contributes appreciably only in the highest layers considered here, where the assumed electron temperature is not far from the solar effective temperature; its maximum contribution is 5 per cent at $\tau_0 = 10^{-5}$. The relationship between P_g and P_e for given T was derived from the ionization tables of Strömgren (1944); since the ionization of hydrogen depends mainly on the unknown intensity of radiation

* These results provide direct evidence for the view that the energy of turbulent motion in the photosphere is mainly dissipated into heat when acoustic or hydromagnetic waves are propagated upwards into the lowest part of the chromosphere (cf. Athay and Thomas, 1956).

in the Lyman continuum, some uncertainty exists about this relationship in the highest layers ($\tau_0 \leq 0{\cdot}002$), but the calculation should be reliable at greater depths where metals supply most of the free electrons. The true value of B is, of course, rather uncertain, and the effects of carbon, nitrogen, and oxygen are neglected.

The model is given in Table II: the quantity η in the third column is the ratio $k_0^{(T)}/k_0$ of equilibrium opacity to actual opacity at $\lambda 5000$. The height scale, H, here represents depth in km below the layer $\tau_0 = 10^{-5}$.

TABLE II

Adopted Model

τ_0	T	η	$\log P_g$	$\log P_e$	$k_0\rho$ in units of 10^{-7} cm^{-1}	H km	τ_{3640}
0·00000							
0·00001	5800	0·7	1·565	−0·655	0·000007	0	0·00008
0·0001	5600	0·7	2·42	−0·445	0·000084	298	0·00040
0·0002	5525	0·75	2·68	−0·39	0·000178	378	0·00060
0·0004	5430	0·8	2·95	−0·355	0·000343	458	0·00092
0·0008	5315	0·85	3·21	−0·33	0·000731	536	0·00144
0·0016	5155	0·9	3·48	−0·33	0·00147	611	0·00226
0·0032	4900	1·0	3·73	−0·28	0·00345	683	0·00360
0·005	4700	1·0	3·87	−0·22	0·00681	721	0·00493
0·008	4430	1·1	4·00	−0·19	0·0121	752	0·00710
0·015	4100	1·2	4·16	−0·11	0·0294	774	0·0121
0·030	4500	1·0	4·32	+0·105	0·0507	812	0·0228
0·060	4800		4·49	0·325	0·0897	856	0·0444
0·085	4880		4·58	0·435	0·130	879	0·0625
0·100	5010		4·62	0·51	0·148	891	0·0734
0·12	5090		4·67	0·585	0·180	903	0·088
0·14	5150		4·705	0·64	0·211	913	0·103
0·16	5230		4·74	0·69	0·234	922	0·117
0·18	5300		4·76	0·76	0·274	930	0·132
0·20	5370		4·79	0·80	0·298	937	0·147
0·25	5520		4·84	0·925	0·388	952	0·186
0·30	5620		4·88	1·01	0·471	964	0·226
0·40	5800		4·93	1·18	0·685	981	0·308
0·50	5910		4·97	1·295	0·885	994	0·393
0·60	6050		5·00	1·43	1·15	1004	0·482
0·80	6250		5·05	1·61	1·67	1018	0·673
1·00	6430		5·08	1·78	2·33	1028	0·881
1·25	6615		5·10	1·945	3·15	1038	1·166
1·50	6780		5·12	2·08	4·04	1045	1·476
1·75	6920		5·14	2·185	4·93	1050	1·811
2·00	7100		5·15	2·33	6·43	1055	2·172

4. *Calculation of emitted intensity in the continuum.*—The geometry of the problem is illustrated in Fig. 2. C is the centre of the Sun and the circular arc represents the layer $\tau_0 = 10^{-5}$, all emission from points above this layer being neglected; x is the distance from this boundary to the point on (or off) the solar disk where the intensity is measured, y the distance from a representative region in the solar atmosphere to the axis OC, and h the depth of this layer below $\tau_0 = 10^{-5}$,

all distances being measured in units of the solar radius R. We then have for the observed emission per cm² per sterad per sec

$$I_\lambda(x) = R\int_{-\infty}^{\infty} \eta\rho k_\lambda(h)B_\lambda(h)e^{-t_\lambda(y)}\,dy$$

$$= R\int_{y_0}^{-y_0} \eta\rho k_\lambda(h)B_\lambda(h)e^{-t_\lambda(y)}(-dy), \tag{9}$$

where

$$t_\lambda(y_1) = R\int_{y_0}^{y_1} \rho k_\lambda(h)(-dy), \tag{10}$$

$$y_0 = (2x - x^2)^{1/2}, \tag{11}$$

and

$$(1-h)^2 = y^2 + (1-x)^2. \tag{12}$$

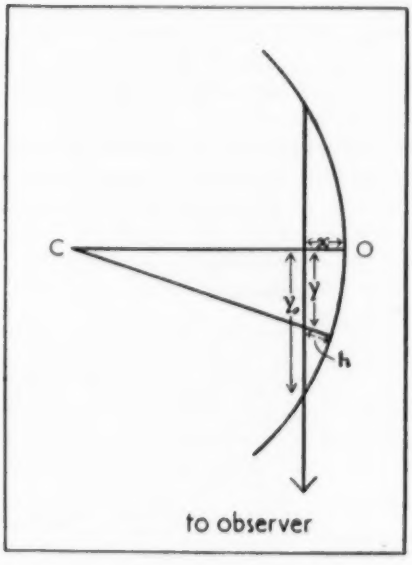


FIG. 2.—*Geometry of emission near the solar limb.*

The numerical integration of equation (9) was carried out with the aid of the EDSAC at the University Mathematical Laboratory, Cambridge. The basic steps in the calculation are the following:

(1) Feed into the machine a table giving selected values of h and the corresponding values of T and of ρk_λ for a given wave-length.

(2) Choose a value of x and calculate the corresponding value of y_0 from equation (11). During the course of the quadratures, the independent variable y is diminished by steps Δy, the choice of which depends on considerations given below. For each value of y, the corresponding value of h follows from (12), but, to save the iteration time involved in the use of a square-root subroutine, (12) was approximated by writing

$$h = \tfrac{1}{2}(2x - x^2 - y^2 + h^2) = \tfrac{1}{2}[2x - x^2 - y^2 + \tfrac{1}{8}(2x - x^2 - y^2)^2]. \tag{13}$$

The error involved in the use of (13) is $\frac{1}{2}h^3$ and thus attains a maximum value of 2×10^{-9} (i.e. 10^{-3} km). The value of h is then employed to find the corresponding values of ρk_λ and of T by Langrangian interpolation.

(3) Evaluate $t_\lambda(y)$ by quadrature on ρk_λ, using (10).

(4) Evaluate $I_\lambda(x)$ by quadrature on (9).

The integration was concluded when $t_\lambda \geqslant 12$ or when $y \leqslant -y_0$, whichever occurred sooner.

The factor $\eta(h)$ was omitted from the machine integration, an estimated average value being applied subsequently as a correction.*

The method of quadrature chosen was Simpson's rule. The reason for this is that it enables the same values of $\rho k_\lambda(y)$ to be used for both t_λ and I_λ, the superior accuracy required in t_λ (which enters as an exponent in (9)) being readily obtained by choosing intervals Δy for t_λ exactly half of those chosen for the quadrature of I_λ. The general attractiveness of other quadrature procedures, e.g. the Gaussian, is removed in a problem of this kind by the necessity of carrying out a considerable number (e.g. 6) of lengthy interpolations for ρk_λ to obtain the appropriate value of t_λ *at each step*, though of course the number of such steps would be greatly diminished.

The choice of interval is governed by the fact that the relative error involved in a Simpson's-rule step for a function of the form

$$f(y) \propto e^{y/a}$$

is

$$\epsilon = \frac{\Delta y^4 f^{IV}(\theta)}{2880 f(\phi)} \simeq \frac{1}{2880}\left(\frac{\Delta y}{a}\right)^4, \qquad (14)$$

where θ, ϕ are values of the independent variable contained within the range of integration Δy. For our calculation, therefore, the requirements are that ρk_λ should not increase by more than a factor e over the half interval Δy and that t_λ should not increase by more than 1 over the whole interval $2\Delta y$. Since ρk_λ actually increases rapidly only in the outer layers where $t_\lambda < 0{\cdot}01$, so that an error of a few per cent in t_λ can be readily tolerated, only the second criterion is of importance. An arbitrary value of Δy was fed in at the beginning of each integration and the programme was arranged to double or halve its value or leave it unchanged after each whole step in such a way as to keep the estimated whole interval $2\Delta t = 2R\rho k_\lambda \Delta y$ between the limits 0·075 and 0·2, where R is the solar radius in cm. For $t_\lambda \geqslant 5$, Δy was doubled automatically if $2R\rho k_\lambda \Delta y < 0{\cdot}7$, since the relative error in the contribution to I_λ from this region is attenuated by a factor $\geqslant 100$ in the total. In this way, the total number of integration steps was kept down to 30, each step taking about 30 seconds.

5. *Results and comparison with observation.*—Calculations of the emergent intensity for various values of x were carried out at $\lambda 4700$, $\lambda 3650$, and $\lambda 3640$, assuming the excitation and ionization of hydrogen to be in accordance with local thermodynamic equilibrium. In other words, η was taken as unity for neutral hydrogen absorption and emission. As regards excitation to the Balmer level, this assumption may be tentatively and qualitatively justified by remarking that the chromosphere is optically thick in Lyman α radiation so that the energy

* Since free-free emission is small compared with bound-free emission for $\lambda < 7000$, η was taken directly from Table I. The uncertainty in η is too large to justify its explicit inclusion in the integration. The adopted effective values are given in Table III.

density at this frequency should be close to that in an enclosure at the local electron temperature; the assumption of equilibrium ionization may be vitiated by the degradation of Lyman continuum radiation through ionization-recombination processes. The equilibrium hypothesis will be re-examined below in the light of observational results.

We assume

$$\frac{a_{4700}(\mathrm{H}^-)}{a_{5000}(\mathrm{H}^-)} = 0{\cdot}936,$$

$$\frac{a_{3650}(\mathrm{H}^-)}{a_{5000}(\mathrm{H}^-)} = 0{\cdot}708,$$

and $$a_{3640}(\mathrm{H}) = 6{\cdot}30 \times 10^{-17-10{\cdot}16\Theta}[1 + 0{\cdot}3 \times 10^{-1{\cdot}88\Theta} + 0{\cdot}1 \times 10^{-2{\cdot}54\Theta}],$$

where $\Theta = 5040/T$. The stimulated emission factor is negligible at these temperatures and wave-lengths.

The results obtained from the adopted model are given in Table III, the limb being taken as the layer where the line-of-sight optical depth t_{4700} reaches the value 1·00. In fact, t_λ reaches this value at $x = 0{\cdot}00104$ (i.e. $\tau_0 = 0{\cdot}005$) for all three wave-lengths considered.

Similar calculations were carried out at $\lambda 4700$ for Böhm's theoretical model photosphere in non-grey radiative equilibrium (Böhm, 1954a) and for the preliminary model illustrated by a broken line in Fig. 1. These results are not tabulated, but they are illustrated in Fig. 3 and they give an estimate of the effect of changing the model. The limb darkening of each model at $\lambda 4700$ was also calculated in the plane-parallel approximation, assuming a linear expression for the Planck function $B_\lambda(\tau_\lambda)$ in the range $\tau_0 > 2{\cdot}0$. For this purpose, use was made of sub-routines in the EDSAC library based on the Gauss–Laguerre 5-point and 6-point formulae (Kopal, 1955); the repetition of the quadratures with a different number of integration points enables their precision to be checked.

The comparison with observation is divided for convenience into two groups: (i) the emission gradient and absolute values in the low chromosphere, and (ii) limb darkening at the extreme solar limb.

(i) ***Emission in the low chromosphere.***—Fig. 3 shows the intensity I_{4700} as calculated from the adopted model, from the preliminary model shown by a broken line in Fig. 1, and from Böhm's radiative-equilibrium model, together with the observational results at neighbouring wave-lengths of Lindblad (1953; quoted by Wildt, 1947), Kristenson (1951, 1953), and of the High Altitude Observatory group as analysed by Athay, Menzel, Pecker and Thomas (1955). A further measurement of the emission gradient, at $\lambda 6000$, by Henriksen (1954) gives a result differing by only 2 per cent from that of Kristenson at $\lambda 4100$. The first two sets of results are on an arbitrary scale and are accordingly shifted vertically to fit the computed curve as well as possible. The last set was derived by graphical differentiation of the total intensities E_λ of vertical strips of the chromosphere given in Table I of the paper by Athay *et al.** The probable error of these results is roughly 2·5 times that of E_λ, i.e. about 0·25 in the logarithm. The curve obtained from the adopted model is in good general

* In general, it is dangerous to differentiate a curve derived from observation. However, the integrated intensity of a strip extending upwards from the lunar limb could be systematically affected below a certain height by a small prominence, the effect of which will be confined to a single height when differences between one strip and another are used.

TABLE III

Calculated continuum intensities and tangential optical depths near the solar limb

I_λ is given in units of 10^{14} erg cm^{-2} sec^{-1} cm^{-1} sterad^{-1}

x	τ_0	Height above limb, km	Height above limb, solar radii	η	λ 4700 t_λ	λ 4700 I_λ	λ 3650 t_λ	λ 3650 I_λ	λ 3640 t_λ	λ 3640 I_λ	$D = \log \frac{I_{3640}}{I_{3650}}$
0·01133		−7150	−0·01029	1·00		1·332		1·100		1·048	−0·02
0·00503		−2780	−0·00399	1·00		1·040		0·780		0·741	−0·02
0·00252		−1030	−0·00148	1·00		0·810		0·544		0·503	−0·02
0·00125	0·0724	− 160	−0·00023	1·00	>12	0·598	10·96	0·315	>12	0·315	0·00
0·00110	0·0123	− 42	−0·00006	1·05	2·75	0·627	2·00	0·310	2·46	0·324	+0·02
0·00100	0·0036	+ 28	+0·00004	1·0	0·69	0·589	0·53	0·306	0·72	0·355	+0·06
0·00095	0·0026	63	0·00009	1·0	0·50	0·513	0·36	0·279	0·52	0·339	+0·08
0·00085	0·0014	132	0·00019	0·9	0·26	0·335	0·19	0·176	0·32	0·257	+0·16
0·00060	0·00028	306	0·00044	0·8	0·059	0·092	0·038	0·044	0·11	0·151	+0·54
0·00040	0·00010	445	0·00064	0·75	0·016	0·027	0·012	0·014	0·058	0·091	+0·81
0·00020	0·00003	585	0·00084	0·7	0·0026	0·005	0·0019	0·0025	0·024	0·044	+1·25
0·00010	0·00002	655	0·00094	0·7	0·0013	0·002	0·0009	0·0012	0·014	0·034	+1·45
			$\frac{t_\lambda}{\tau_\lambda}$ (mean value)		208		200		180		

agreement with the observational results; the model leads to a very slight limb brightening which, however, is insufficient to reproduce the very intense brightening (actually a doubling of intensity between −120 km and +45 km) derived from the measurements of the High Altitude Observatory group. A higher surface temperature would reproduce this strong brightening, but at the expense of appreciable departures from the Lindblad–Kristenson gradient. Pending further observations, it would seem that the present model gives the best available overall agreement with the various sets of measurements that have been carried out. In particular, the observation of limb brightening (confirmation of which by further measurements would of course be very desirable) is a direct confirmation of the existence of a minimum of electron temperature in the photosphere; furthermore, the absolute value of I_{4700} derived from the Harvard measurement at −120 km agrees well with the assumption that $T = 4100$ deg. at this minimum (as inferred from excitation temperatures and from centre-limb variations in the infra-red Fe I lines), though it could also be reconciled with a higher value.

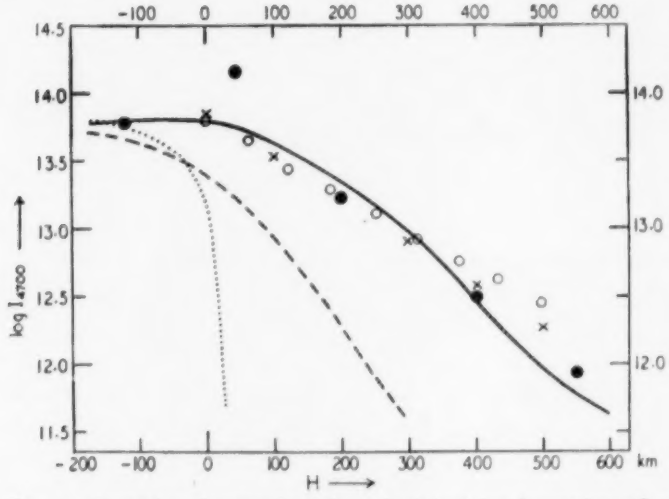


FIG. 3.—*Comparison of computed and observed emission in low chromosphere at λ 4700.*

Open circles:	*Lindblad.*	———	*Suggested model.*
Crosses:	*Kristenson.*	— — —	*Preliminary model.*
Filled circles:	*Athay et al.*	.....	*Radiative equilibrium model (Böhm).*

In Fig. 4, the calculated free–bound emission at the head of the Balmer continuum, λ3640, is compared with the observations of the High Altitude Observatory group. The computed curve is obtained from Table III by forming the difference $I_{3640} - I_{3650}$, and it agrees in its general behaviour with the results of Athay *et al.* (1955), particularly in reaching a maximum a few hundred km above the limb. The fact that the computed curve falls below the observations above 400 km is merely due to the fact that higher layers (above 725 km) were neglected in the computations. Below 400 km, the computed difference exceeds the measured value by as much as a factor 3, but it is doubtful whether this discrepancy is significant. In view of the many uncertainties involved, it is not

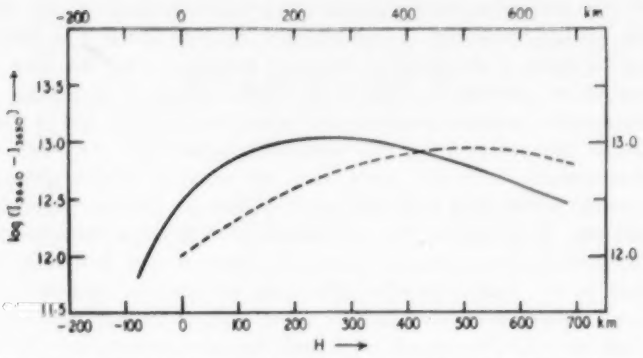


FIG. 4.—*Comparison of computed and observed intensities in Balmer continuum.*
Full line: adopted model.
Broken line: Athay et al., "Best fit".

worth trying to make the small alterations in the model needed to remove it (a change in η would, for example, be sufficient). The fact that the computed value of $I_{3640} - I_{3650}$ agrees in sign and order of magnitude with the observed values confirms our assumption of approximate thermodynamic equilibrium in the excitation and ionization of hydrogen in those layers where self-absorption is significant. For, integrating equation (9) for an isothermal layer with constant η,

$$I_\lambda(x) = \eta B_\lambda(T)(1 - e^{-t_\lambda \eta}), \tag{15}$$

where t_λ is calculated assuming thermodynamic equilibrium; we assume that electron, proton, and neutral hydrogen densities are kept fixed. Now, at +63 km for example, a non-equilibrium factor $b_2 = 5$ for the Balmer level corresponds to $\eta = 1{\cdot}0$ at λ 3650 and $\eta = 0{\cdot}45$ at λ 3640 (Table III). Hence I_{3640} would be reduced from $0{\cdot}339 \times 10^{14}$ to $0{\cdot}264 \times 10^{14}$, and the difference $I_{3640} - I_{3650}$ would come out negative.

It is interesting to compare our result $b_2 \simeq 1$ with the large values obtained by Pecker and Pecker (1956) for the higher b_n's at greater heights (at and above 1000 km). It remains to be seen whether a continuous gradation in b_2 with increasing height can be established.

In the higher regions considered here, self-absorption becomes unimportant so that the presence or absence of equilibrium excitation cannot be checked by means of these results. At 400 km, as might be expected, our calculated value of the Balmer discontinuity $D = \log(I_{3640}/I_{3650})$ falls well within the range of measurements obtained by Koelbloed and Veltman (1951). The results thus suggest that our assumption of equilibrium excitation to the Balmer level near the limb was justified at least within a factor of 2 or 3, and that the electron densities deduced from the assumption of equilibrium ionization are also nearly correct.

(ii) *Darkening at the extreme solar limb.*—The calculated limb-darkening curve for λ 4700 at the extreme limb of the Sun is given in Fig. 5, together with several observational results. The continuous lines represent the limb darkening derived from our adopted model (thick line) and from the radiative-equilibrium model (Böhm, 1954a; thin line) taking curvature into account, while the broken lines

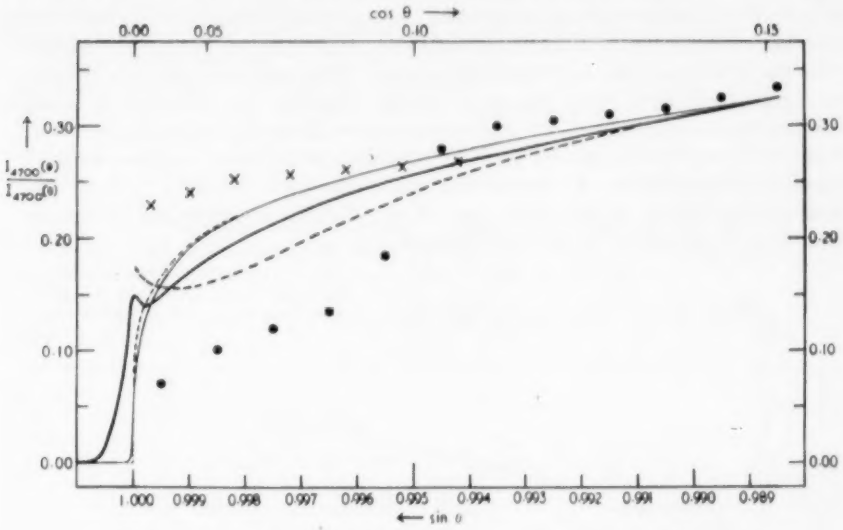


FIG. 5.—*Computed and observed darkening at extreme solar limb for* λ4700.
Full lines: darkening computed allowing for curvature. *Filled circles: Heyden.*
Broken lines: darkening computed in plane-parallel approx. *Crosses: Wesselink.*

represent the plane-parallel approximation. In Böhm's model, curvature makes only a very slight difference even at the extreme limb; in ours, its effect is appreciable for $\sin\theta \geqslant 0{\cdot}992$ (i.e. $\cos\theta \leqslant 0{\cdot}13$). The relevant observational results are those of Heyden *et al.* (1953), whose measurements at λλ 4800, 4400 have been interpolated for λ 4700, and of Wesselink *et al.* (1940) at an effective wave-length λ 4540. For $\sin\theta \leqslant 0{\cdot}995$ ($\cos\theta \geqslant 0{\cdot}10$), the computed curves for both models agree well with the observed trend (the intensity *scale* inevitably involves some degree of arbitrariness), Böhm's model being slightly superior to ours; closer to the limb, a marked divergence appears between all sets of results. A detailed comparison of the observational results at various wave-lengths is given in Heyden's paper. In view of the analysis by Fellgett and Schmeidler (1952), one may conclude that the large discrepancies between different sets of observational results close to the limb are due to the magnification of small errors in the original observations during the reduction, which involves both differencing and "unsmearing" the derived intensity of a given crescent to obtain $I_\lambda(\theta)$. In spite of these difficulties (and thin clouds at the Dutch 1936 eclipse expedition), the trend obtained by Wesselink agrees remarkably well with that derived in a totally different manner by ten Bruggencate, Gollnow and Jäger (1950; as rediscussed by Fellgett and Schmeidler, 1952) at λ 6300; whereas the scatter in the raw data of Heyden *et al.* casts some doubt on the detailed trend of $I_\lambda(\theta)$ derived by them close to the limb. The solution to these difficulties can be obtained by eliminating the reduction process altogether and proceeding directly from a given model atmosphere to the raw observational data. This procedure has been carried out by Hubenet and de Jager (1956); their results confirm this criticism and also appear to lead to a model somewhat similar to ours for $\tau_0 \geqslant 0{\cdot}01$.

The last column of Table III gives calculated values of the Balmer discontinuity close to the solar limb. By calculating the limb darkening for λλ 3640,

3650 in the plane-parallel approximation, using linear B_λ ($\tau_0 > 2{\cdot}0$), these values have been extended to the centre of the disk. Fig. 6 shows the result, together with the measurements of Peyturaux (1954). The agreement is very good in the range $\cos\theta \leqslant 0{\cdot}5$; it is less good at the centre of the disk, but it is still satisfactory there in view of the approximations made in the computation. This result confirms the calculated electron pressures for $\tau_0 \geqslant 0{\cdot}1$ and also suggests equilibrium excitation to the second quantum level in these layers. An observational study of the behaviour of the Balmer discontinuity across the limb itself would be of the greatest interest.

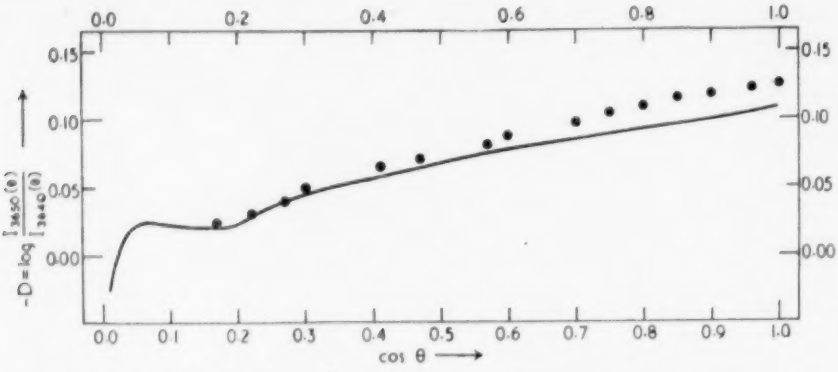


FIG. 6.—*Calculated and measured centre-limb variation in Balmer discontinuity.*
Full line: adopted model.
Filled circles: measured values (Peyturaux, 1954).

I am indebted to Dr R. Grant Athay for his very helpful criticism, to Dr M. V. Wilkes for his permission to use the EDSAC for this investigation, and to the Master and Fellows of Sidney Sussex College, Cambridge, for a Research Fellowship.

Royal Greenwich Observatory,
Herstmonceux Castle,
Hailsham, Sussex:
1956 *November.*

References

Aller, L. H., 1953, *Astrophysics*, Vol. I (New York : Ronald Press Co.), p. 226.
Athay, R. G. and Menzel, D. H., 1956, *Ap. J.*, **123**, 285.
Athay, R. G., Menzel, D. H., Pecker, J. C. and Thomas, R. N., 1955, *Ap. J. Supp.*, **1**, 505.
Athay, R. G. and Thomas, R. N., 1956, *Ap. J.*, **123**, 309.
Bhatnagar, P. L., Krook, M., Menzel, D. H. and Thomas, R. N., 1955, *Vistas in Astronomy*, Vol. I (London : Pergamon Press), p. 296.
Böhm, K. H., 1954 a, *Zs. f. Ap.*, **34**, 182.
Böhm, K. H., 1954 b, *Zs. f. Ap.*, **35**, 179.
Böhm-Vitense, E., 1954, *Zs. f. Ap.*, **34**, 209.
Böhm-Vitense, E., 1955, *Zs. f. Ap.*, **36**, 145.
Bruggencate, P. ten, Gollnow, H. and Jäger, F. W., 1950, *Zs. f. Ap.*, **27**, 223.
Chandrasekhar, S., 1945, *Ap. J.*, **102**, 395.
Chandrasekhar, S. and Breen, F. H., 1946, *Ap. J.*, **104**, 430.
Fellgett, P. B. and Schmeidler, F. B., 1952, *M.N.*, **112**, 445.

Henriksen, S. W., 1954, *Ap. J.*, **120**, 521.
Heyden, F. J., Beck, C. A., Lovell, D. J., Fussell, W. B., Hancock, J. H. and Hulburt, E. O., 1953, *Ap. J.*, **118**, 412.
Hubenet, H. and Jager, C. de, 1956, *B.A.N.*, **13**, 43.
Jager, C. de, 1952, *Rech. Astr. de l'Obs. d'Utrecht*, **13**, 1.
Jager, C. de, 1954, *Nature*, **173**, 680.
Jager, C. de, 1955, *Ann. de Geophys.*, **11**, 330.
Koelbloed, D. and Veltman, W., 1951, *Proc. Kon. Ned. Ak. v. Wetensch*, B, **54**, 25; Amsterdam Circ. No. 3.
Kopal, Z., 1955, *Numerical Analysis* (London : Chapman and Hall), pp. 370, 525.
Kristenson, H., 1951, *Stockholms Obs. Ann.*, **17**, No. 1.
Lindblad, B. and Kristenson, H., 1953, *Atti del Convegno Volta*, p. 61.
McCrea, W. H., 1929, *M.N.*, **89**, 718.
Menzel, D. H., Baker, J. G. and Goldberg, L., 1938, *Ap. J.*, **87**, 81.
Milne, W. E., 1949, *Numerical Calculus* (Princeton : at the University Press), p. 72.
Minnaert, M., 1953, *The Sun*, ed. G. P. Kuiper (Chicago : at the University Press), p. 164.
Pagel, B. E. J., 1955, *M.N.*, **115**, 493.
Pagel, B. E. J., 1957, *Ap. J.*, **125**, 298.
Pecker, C. and Pecker, J. C., 1956, *C.R.*, **242**, 994; Cont. I.A.P., ser. A, No. 213.
Peyturaux, R., 1954, *C.R.*, **239**, 1459; Cont. I.A.P., ser. A, No. 176.
Plaskett, H. H., 1955, *M.N.*, **115**, 256.
Redman, R. O. and Suemoto, Z., 1954, *M.N.*, **114**, 524.
Regemorter, H. van, 1953, *C.R.*, **236**, 1536; Cont. I.A.P., ser. A, No. 145.
Strömgren, B., 1944, *Pub. og mind. Medd. Københavns Obs.*, No. 138.
Thomson, J. J., 1912, *Phil. Mag.*, **23**, 419.
Unsöld, A., 1953, *Atti del Convegno Volta*, p. 27.
Unsöld, A., 1955, *Physik der Sternatmosphären*, 2nd ed. (Berlin : Julius Springer), p. 373.
Voigt, H. H., 1956, *Zs. f. Ap.*, **40**, 157.
Wesselink, A. J., Ferwerda, J. G. and Uitterdijk, J., 1940, *B.A.N.*, **9**, 81.
Wildt, R., 1947, *Ap. J.*, **105**, 36.
Woolley, R. v. d. R., 1934, *M.N.*, **95**, 101.
Woolley, R. v. d. R. and Stibbs, D. W. N., 1953, *The Outer Layers of a Star* (Oxford : at the Clarendon Press), p. 57.
Zanstra, H., 1950, *Proc. Kon. Ned. Ak. v. Wetensch.* B, **53**, 1289; Amsterdam Circ., No. 1.

A STUDY OF A BLOW-OFF PROMINENCE (1956 MAY 18)

M. A. Ellison

(Received 1956 November 21)

Summary

Observational data relating to a blow-off prominence have been collected with the combined spectrohelioscope and spectrograph and are discussed. Photometric measures of Hα brightness, in terms of the centre of disk continuum, were made from two spectra. The importance of recording blow-off phenomena during the International Geophysical Year is stressed.

---

1. *Introduction.*—Typical quiescent prominences, observed as dark filaments on the solar disk, are among the most stable of all solar features. M. and L. d'Azambuja, in their notable work on the life-histories of the filaments (**1**), give the average duration as three to four solar rotations. In exceptional cases they may persist for more than six rotations.

However, continuous observations made with the spectrohelioscope show that the normal quiescent life of a filament is frequently interrupted by sporadic outbreaks of internal activity. At these times the higher parts of the prominence develop into a complex pattern of streamers which show sightline components of velocity of the order of ±30 km/sec. Hunting backwards and forwards over a range of 1–2 A with the line-shifter, one observes rapid changes both of shape and luminosity between adjacent wave-lengths. The prominence "wriggles" as the line-shifter is moved. These periods of activation may last for several hours, after which, in many cases, the internal turbulence dies away and the prominence resumes its quiescent state. In this more normal condition the luminosity fades out uniformly at equal settings of the line-shifter on either side of the line centre.

Less frequently this state of internal activation develops into the violent "blow-off" phase affecting the whole, or a part, of the filament. The central portion of the filament begins to ascend, slowly at first but with increasing velocity, while one or both ends remain anchored to the chromosphere. An expanding arch is formed, the top of which may rise to heights of 500×10^3 km before it becomes too faint to follow further ; by this time the outward velocity along the solar radius will have reached a value of the order of 500 km/sec.

Many examples of the blow-off phenomenon have been described by the d'Azambujas (**1**), who refer to them as "disparitions brusques". They agree with Waldmeier (**2**) in stating that the disparition brusque is a normal phase in the evolution of the filaments. Equally remarkable is their conclusion that only one blow-off in three leads to the final dissolution of a filament. In most cases the blow-off phase is followed during a period of several days by the gradual *re-growth* of the prominence within the lower corona, so that it eventually re-appears with almost its original aspect and size. There is no observational evidence for the return of the blown off material from outer space, and the possibility of such a return to precisely the original site seems altogether too improbable.

We must therefore envisage the situation that what survives the blow-off catastrophe is not the material of the original filament but rather the "field" conditions within the lower corona which caused its generation and enabled it to radiate in Hα, K_3 etc. Whatever these field conditions may be, and however they are maintained, it is evidently their persistence and survival which leads to the peculiar stability and re-formation of the filaments.

2. *The blow-off prominence of* 1956 *May* 18.—The great centres of sunspot activity in the northern hemisphere which occurred during the months of January–March 1956 led to a steep rise in the formation of filaments in latitudes $+30°$ to $+60°$. These reached their greatest development during May.

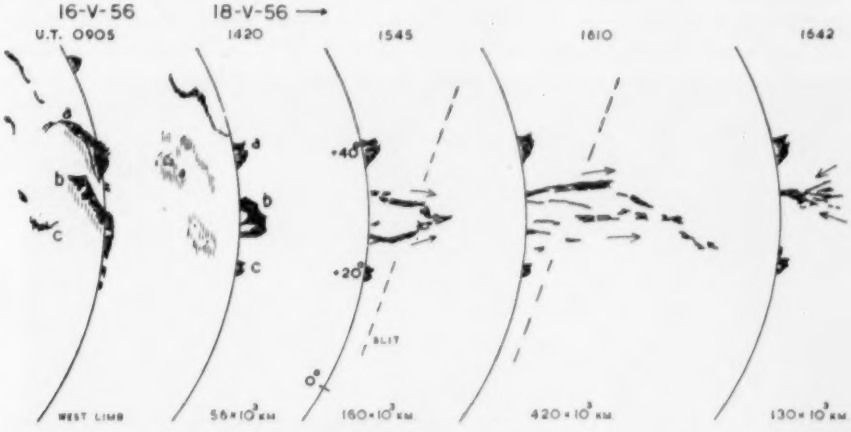


FIG. 1.—*Quiescent filaments on the W limb,* 1956 *May* 16, *and the blow-off stages of filament b on May* 18. *The dashed line shows the orientation of the spectrograph slit when the Hα spectra were taken at* $16^h\ 01^m$ *and* $16^h\ 04^m$.

Three of these filaments (Fig. 1, *a*, *b*, *c*) were approaching the west limb on May 16, *a* and *b* having been followed right across the disk. By May 18 they were visible as three prominences seen in elevation at the limb. Prominence *b* (extending from $+26°$ to $+33°$) was contained in a plane inclined at about 30° to the line of sight with its northern end nearest to the observer; its total length at this time, measured along a great circle, was close to 400×10^3 km. At $14^h\ 20^m$ U.T. on the 18th *b* was seen to be active and the "trunks" attaching it to the chromosphere were already lengthening. By $15^h\ 45^m$ the top of the arch had risen to a height of 160×10^3 km and the two legs had been drawn out into streamlined structures of bright knots. At this time the line-shifter of the spectrohelioscope showed that the southern leg had a sightline velocity component of $+100$ km/sec (recession) and the northern leg -95 km/sec (approach). From what has just been said about the orientation of the filament we interpret the measures at this time as representing an arch which was expanding with about equal velocities sideways and upwards in its own plane, but with its two feet still anchored in the chromosphere.

At $16^h\ 10^m$ the top of the arch had reached a height of 420×10^3 km above the limb and could not be followed further owing to its increasing faintness. By $16^h\ 42^m$ all that could be seen visually were portions of the northern leg

of the arch which were now streaming inwards towards a point in the chromosphere at $+34°$.

A noteworthy feature throughout was the immobility of the prominences *a* and *c*, which were located only some 120×10^3 km to north and south of all this intense activity.

3. *Photometry.*—Calibrated plates of the Hα line were taken at $16^h\,01^m$ and $16^h\,04^m$ with the slit of the spectrograph crossing the arch at a mean height of 88×10^3 km above the limb (Fig. 1). Plate 8 shows Hα as it was recorded at $16^h\,04^m$, the line being then highly distorted by the sightline components of velocity within the legs of the arch. These plates were calibrated immediately afterwards in the spectral light from the centre of the solar disk, so that we have the means of deducing the Hα brightness of the various points (Plate 8, A, B, C, etc.) of the prominence which were traversed by the slit. Our standard methods of photometry were used, as described in an earlier publication (**3**), and the figures given in Table I are expressed as fractions of the centre of the disk continuum at a wave-length 15 A outside the Fraunhofer Hα line.

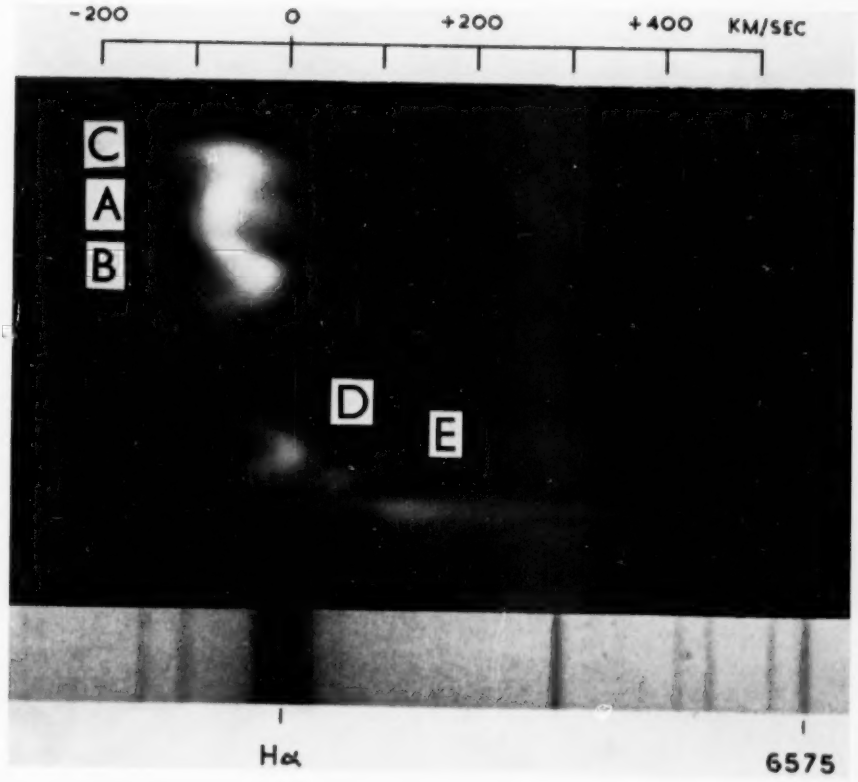
TABLE I

Plate	Region				
	A	B	C	D	E
$16^h\,01^m$	0·050	0·040	0·042	...	...
$16^h\,04^m$	0·047	0·037	0·033	0·006	0·005
Height above limb (10^3 km)	98	...	...	...	77

The values derived from the two plates are thus in close agreement, despite the rapid movements which were taking place within the prominence at this time. Our previous measures (**3**) have given the value 0·062 for the mean central intensity of the Hα line of quiescent limb prominences at a mean height of 33×10^3 km above the limb. Evidently the luminous matter in the blow-off filament had a comparable brightness (0·05) at three times this height, or 98×10^3 km.

4. *Doppler shifts.*—The sightline components of velocity can be seen from Plate 8 (taken at $16^h\,04^m$ U.T.) to range from -120 km/sec in the northern leg (C) of the arch to $+360$ km/sec in the outermost portion (E) of the southern leg. These are in the same sense as, though greater than, the measures made with the spectrohelioscope line-shifter 19 minutes earlier. It is also interesting to note how little of this luminous material coincides in wave-length with the undisplaced Hα line. Thus a spectroheliograph, or even a filter accepting a passband of 0·7A, would have recorded only about one-fifth part of the prominence which appears on the spectrogram.

5. *Height–time diagram.*—In Fig. 2 we have a height-time plot for the topmost part of the arch. This was constructed from three measures of height above the solar limb made at Edinburgh (○) and three made at Meudon (•); the latter were kindly supplied by L. d'Azambuja from spectroheliograms. Although the points are not sufficiently numerous to give very accurate determinations of velocity, they clearly indicate the rapid increase of the velocity component directed along the solar radius. In the early stages before $15^h\,00^m$ the slope of the diagram indicates an outward velocity of 20 km/sec. After $16^h\,07^m$ the velocity exceeds 600 km/sec. The observed heights in Fig. 2 have been joined



The Hα *line of the blow-off prominence distorted by Doppler shifts, 1956 May 18, 16h 04m U.T. The lettering refers to Table I.*

by straight lines, but this is not meant to imply that the velocity along the solar radius assumed a series of constant values.

6. *Conclusions.*—The observations in the present case suggest the concept of a uniformly expanding arch in the early stages of the blow-off, followed by the rapid outward acceleration of the top. This last phase is accompanied by the breaking of the legs, the lower portion of the northern leg returning to the chromosphere. In the previous literature we find a number of well observed examples which appear to have followed a similar sequence. Detailed studies have been made of the following : 1919 May 29 and July 15 (Pettit (**4**)), in which the plane of the filaments was nearly at right angles to the line of sight ; 1937 September 2 (Waldmeier (**2**)) ; 1946 June 4 (Pettit (**5**) and Roberts (**6**)) ; 1947 June 11 (Ellison (**7**)). Olivieri has reported two interesting examples taken from Meudon spectroheliograms, the one on 1948 September 7–8 (**8**) and the other on 1956 February 9 (**9**). In the first of these, an Hα picture taken four days later shows

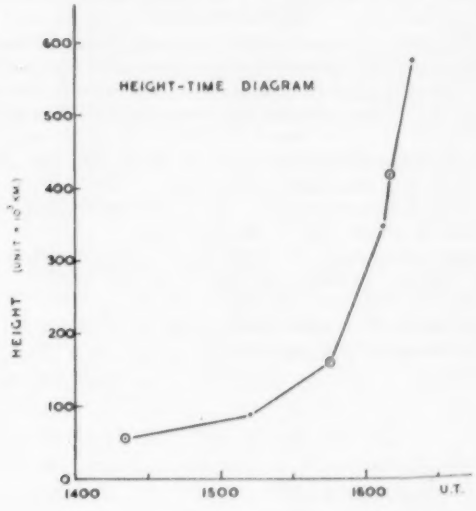


FIG. 2.—*Height–time diagram for the highest point of the arch. Edinburgh measures* (○); *Meudon measures* (•).

the filament in the process of re-forming. While there is no evidence that the blow-off phase of a quiescent filament is initiated by flare activity, or indeed by any other visible chromospheric phenomenon, this question merits further examination when continuous patrol films of the chromosphere become available.

Such high-speed projection of prominence material through the corona may well be associated with the generation of radio waves of frequencies appropriate to the various levels. There is also the possibility of geomagnetic effects occurring on the Earth after a suitable travel time in those cases where the blow-off has been observed near the centre of the solar disk. For these and similar reasons it is desirable that all examples of the blow-off of filaments should be recorded and classified at solar observatories during the I.G.Y.

Ohman (private communication) has proposed the following simple scheme of classification for inclusion in the I.G.Y. *Solar Patrol Forms*:

	Importance Class	Apparent Length
Blow-off filaments observed either on the disk or at the limb	1	10–20
	2	20–40
	3	40–80
	3+	>80

The importance is to be determined by the apparent length of the moving filament (expressed as a percentage of the solar radius) according to the figures given in the last column. It is hoped that this method of classification will be generally adopted. A spectrohelioscope provided with a graticule or micrometer at the focus is ideal for this type of observation.

References

(1) M. et Mme L. d'Azambuja, *Étude d'ensemble des protubérances solaires et de leur évolution, Ann. de l'Obs. de Paris*, Section d'Astrophysique a Meudon, VI, 1948, p. 104. See also L. d'Azambuja: *Protubérances solaires a évolution rapide, l'Astronomie*, May 1925, p. 209; *Deux Protubérances remarkables, l'Astronomie*, February 1926, p. 64.
(2) M. Waldmeier, *Aufsteigende Protuberanzen, Zs. f. Ap.*, **15**, 299, 1938.
(3) M. A. Ellison, *Publ. Roy. Obs. Edin.*, **1**, 75, 1952.
(4) E. Pettit, *The forms and motions of the solar prominences, Publ. Yerkes Obs.*, **3**, 205, 1925. See also E. Petitt, *Ap. J.*, **50**, 206, 1919.
(5) E. Pettit, *An eruptive prominence of record height, P.A.S.P.*, **58**, 310, 1946.
(6) W. O. Roberts, *Sky and Telescope*, **5**, Nos. 10 and 11, 1946.
(7) M. A. Ellison, *Jour. B.A.A.*, **57**, 229, 1947.
(8) G. Olivieri, *l'Astronomie*, December 1948, p. 343.
(9) G. Olivieri, *l'Astronomie*, March 1956, p. 120.

THE D_3 EMISSION OF PROMINENCES

J. T. Jefferies

(Communicated by R. G. Giovanelli)

(Received 1956 April 30)

Summary

The computed emission of Hα and D_3 from an isothermal slab model prominence is shown to give agreement with observational data of Brück and Moss when the model has helium abundance one fifth that of hyd gen, thickness $< 2 \times 10^4$ km, temperature between $1{\cdot}1 \times 10^4$ deg. K and $2{\cdot}0 \times 10^4$ deg. K and electron concentration between 10^{10} cm^{-3} to 5×10^{10} cm^{-3}. Some evidence has been found that the higher temperatures are generally associated with lower electron concentrations. Comparison of the present computations with earlier results obtained by the author from an analysis of the contour of the Hα emission of prominences shows that the ratio of the abundances of helium and hydrogen atoms in solar prominences lies in the range 0·12 to 0·4. As for Hα, the self emission in D_3 is generally small compared with the diffusely reflected and transmitted component.

---

1. *Introduction.*—In a recent publication (**1**) the author has analysed observational data of Conway (**2**) and Ellison (**3**) on the central intensities and half-widths of the Hα emission lines of stable prominences in terms of a simple prominence model. It was shown there that the observational results were compatible with models having a thickness in the observed range, kinetic temperature T between 10^4 deg. K and 2×10^4 deg. K, and an electron concentration N_e between 10^{10} cm^{-3} and 5×10^{10} cm^{-3}.

To compute the Hα contours on which the analysis of (**1**) was based, expressions were required for the excited state populations of hydrogen as a function of electron temperature and concentration. The author has recently (**4**) computed the corresponding data for helium, and an analysis similar to (**1**) could now be made for the $D_3(3^3D\text{–}2^3P)$ emission line. In general, however, observations of the D_3 emission have not been made in the necessary detail, the most extensive set of data being that of Brück and Moss (**5, 6, 7**), who made simultaneous measurements of the equivalent widths of Hα and D_3 at points in a considerable number of prominences. In this paper we compute the Hα and D_3 emission from isothermal slab model prominences having various physical properties. Comparison of these results with the observations of Brück and Moss gives information on the range of electron temperature and concentrations of solar prominences.

2. *The equivalent width of Hα.*—As shown in (**1**) the monochromatic specific intensity $I_\nu(0, 1)$ of the Hα radiation emerging at right angles to the plane of a model prominence is easily obtained by means of an Eddington approximation. The value of $I_\nu(0, 1)$ depends on the type of scattering adopted; and in (**1**) two cases were considered, namely a purely coherent and a partially coherent scattering

mechanism. In the latter case—which we adopt here as being theoretically sounder—the source function $\mathfrak{S}_\nu$ for the Hα radiation is written, following Woolley and Stibbs (**8**), in the form

$$\mathfrak{S}_\nu = \frac{1-\lambda}{4\pi}\left[aJ_\nu + bJ_{\nu_0}\right] + \epsilon_\nu/\kappa_\nu, \tag{1}$$

where λ is a parameter defined so that $1-\lambda$ is the probability that an absorbed quantum will be subsequently re-emitted in the same spectral line, ϵ_ν and κ_ν are respectively the monochromatic emission and extinction coefficients. The total intensities J_{ν_0} and J_ν at the centre of the line and at a frequency ν respectively are defined by

$$J_\nu = \int_{4\pi} I_\nu \, d\omega.$$

The constants a and b are defined by

$$a = \frac{\delta_k}{\delta_j + \delta_k}; \qquad b = \frac{\delta_j}{\delta_j + \delta_k},$$

k and j referring respectively to the upper and lower states involved in the transition and δ_j and δ_k being given by

$$\delta_j = \sum_{i<j} A_{ji}/4\pi; \qquad \delta_k = \sum_{i<k} A_{ki}/4\pi,$$

the A's being spontaneous transition rates.

The emergent specific intensity consists of a self-emitted component and one due to the radiation from the solar disk diffusely transmitted and reflected by the slab. The magnitude of this latter component depends on the law of darkening assumed for the disk radiation and we adopt, as in (**1**), a law of the form

$$I(\mu) = 0{\cdot}4I(1)[1 + 1{\cdot}5\mu] \tag{2}$$

where $\mu = \cos\theta$, θ being the angle between the normal to the solar surface and the direction of emergence of the radiation.

With the non-coherent scattering mechanism implied by (1), the intensity of radiation from the prominence at frequency ν may be expressed in terms of the disk intensities at the line centre and at a frequency ν. An approximate expression for the specific intensity $I_\nu(0, 1)$ of the normally emergent radiation can readily be obtained, as shown in (**1**), in the form

$$I_\nu(0, 1) = \frac{\epsilon}{\kappa\lambda} A_\nu(\lambda, \tau_1) + I_{\nu_0} B_\nu(\lambda, \tau_1) + I_\nu C_\nu(\lambda, \tau_1) \tag{3}$$

where τ_1 is the optical thickness of the model at frequency ν and A_ν, B_ν and C_ν are rather lengthy expressions given explicitly in (**1**). The specific intensity I_ν of the normally emergent disk radiation will vary over an absorption line such as Hα; for D_3 however, which does not appear in the Fraunhofer spectrum, it is independent of frequency over the line. In (3) the first term represents the self-emission while the second and third are respectively the non-coherent and the coherent components of the diffusely scattered radiation.

For given electron concentrations and temperatures, expressions for ϵ/κ and λ may be obtained in terms of the transition rates by general methods given by Giovanelli and Jefferies (**9**). To obtain the central intensities and half-widths on which the analysis in (**1**) was based, the contour of the line was computed from (3) for models of given optical thickness, electron concentration and

temperature. For the present work, areas under these profile curves have been measured and the equivalent width of Hα obtained on dividing these by the intensity in one angstrom of the central disk continuum near Hα, which, according to figures given by Munch (**10**), corresponds to that from a black body at 6150 deg. K.

3. *The equivalent width of D_3.*—The emission in the helium D_3 line depends on the electron concentration and temperature, the thickness and the helium abundance. In the absence of better knowledge, we shall suppose helium to be one-fifth as abundant as hydrogen. As we shall be interested only in temperatures above 10^4 deg. K, where hydrogen is largely ionized, the hydrogen concentration will be effectively equal to the electron concentration.

In most cases of interest to us here, the optical depth in D_3 is small, and in this case the specific intensity $I_\nu(0, 1)$ given by (3) can be shown to be of the form

$$I_\nu(0, 1) = \left[\frac{\epsilon}{\kappa} + \frac{1-\lambda}{2}\,\delta I_{\nu_0}\right]\tau_1 \qquad (4)$$

where δ is a factor depending on the law of darkening adopted and in our case is 0·678. Values of ϵ/κ and λ, which are effectively frequency-independent, may be obtained as for Hα while, from standard expressions for the Doppler broadened extinction coefficient κ, we find

$$\tau_1(D_3) = 6{\cdot}3 \times 10^{-11} N_{2^3} z T^{-1/2} \exp\left[-\left(\frac{c\,\triangle\nu}{v\nu}\right)^2\right] \qquad (5)$$

where N_{2^3} is the concentration, in cm^{-3}, of the 2^3 state, z is the thickness of the model in cm, $\Delta\nu = \nu - \nu_0$ and the mean atomic velocity $v = \sqrt{(2kT/M_{H^+})}$. Expressions for N_{2^3} are given in (**4**).

Having computed $I_\nu(0, 1)$ from (3) or (4), integration over the spectral line gives the total D_3 emission. This is reduced to an equivalent width by dividing by the normally emergent continuum intensity in one angstrom near D_3, which corresponds to that from a black body at 6270 deg. K (**10**).

A difficulty arises, however, in that N_1, and so N_{2^3} and $\tau_1(D_3)$, depends quite strongly on the intensity of the ultra-violet continuum of helium and thus on the opacity of the prominence to this radiation. On the assumption that external irradiation of the prominence is negligible in this spectral region, equivalent widths of D_3 have been computed for the limiting cases obtained by assuming either very high or very low opacities, providing limits between which the true result must lie. The range of applicability of each assumption may be checked using expressions for the opacity in terms of N_1 and z (**4**); for $N_e \sim 10^{10}$ cm^{-3} and reasonable thicknesses, the opacity is certainly high when $T \lesssim 10^4$ deg. K and low for $T \gtrsim 2{\cdot}5 \times 10^4$ deg. K. For transitional optical thicknesses $5 \gtrsim \sqrt{3}\lambda\tau_1 \lesssim 0{\cdot}2$ approximate results may be obtained by interpolation between the limiting cases. This can give no more than a rough estimate, however, as the D_3 emission changes by a large factor between these limiting cases. Thus for $T = 1{\cdot}5 \times 10^4$ deg. K, $N_e = 5 \times 10^{10}$ cm^{-3}, the transition occurs when the Hα equivalent width is ~50 mA, the D_3 widths found on the assumptions of low and high opacities then being ~1 mA and ~60 mA respectively. It is clear that the slope of an Hα versus D_3 equivalent width curve will be very small in the transition region but its exact form is difficult to determine.

4. *Comparison with observation.*—The relationships between the computed equivalent widths of Hα and D_3 are shown in Fig. 1 for various values of N_e and T and for an abundance ratio He/H = 0·2. Each point on a curve corresponds to a definite model thickness and, where applicable in the region graphed, the curves have been terminated for a thickness of 2×10^4 km. As this is a generous upper limit for the thickness of a solar prominence, parts of the computed curves lying beyond the terminal point—even if giving agreement with observational data on equivalent widths—cannot correspond to a valid prominence model.

For the $T = 10^4$ deg. K curves, the opacity in the ultra-violet continuum of helium is high except for negligibly small values of the D_3 equivalent width, while for $T = 2{\cdot}5 \times 10^4$ deg. K the opacity is very small over the range of equivalent widths covered. The results obtained for $T = 1{\cdot}5 \times 10^4$ deg. K, $N_e = 5 \times 10^{10}$ cm^{-3} assuming low opacity in the ultra-violet helium continuum are shown by a continuous curve in Fig. 1. For these physical conditions, however, the opacity in the ultra-violet continuum becomes $\sim$1 when the Hα equivalent width is $\sim$50 mA and consequently near this point the true curve will depart from the low opacity curve and will rise monotonically to join the curve obtained by assuming high opacity in the ultra-violet continuum. As this latter curve could not be shown adequately with the scale used in Fig. 1, the coordinates are given of a point near the end of the transitional region. The estimated form of the curve in the transitional region is indicated by a broken line.

For D_3 the emission is almost entirely due to reflection and transmission of the disk radiation, the self-emitted component amounting at most to no more than $\sim$10 per cent of the total. It is of interest to note that, in most cases, a similar conclusion applied for the Hα radiation from prominences (**1**).

Observational points given in the first two publications of Brück and Moss (**5, 6**) are also plotted in Fig. 1. Points for some particular prominences are indicated separately in order to show the general form of the relation between the intensities for a single prominence.

Some general conclusions can be drawn at once. We note first that a reduction of the relative abundance of He would reduce the D_3 emission for a given Hα emission and so make the curves steeper. As it is unlikely that the relative helium abundance is much greater than that adopted, it seems that a temperature of 10^4 deg. K is too low to account for any significant number of prominence observations. This applies for any value of N_e. It appears also that, for any T, N_e must be greater than 10^{10} cm^{-3} since a model with this electron concentration gives observed Hα and D_3 emissions only for excessive thicknesses if at all.

To assess the validity of the higher temperature models, we show in Fig. 2 the relationship between the computed D_3 intensity and model thickness. The broken lines again represent the estimated shapes of the curves in the transition regions. Since the majority of observations correspond to D_3 equivalent widths <6 mA and the thickness could hardly be greater than 2×10^4 km there is a strong indication, when Figs. 1 and 2 are taken together, that few if any observations correspond to models with $T \lesssim 1{\cdot}1 \times 10^4$ deg. K. An upper limit is probably 2×10^4 deg. K but here the unknown helium abundance makes estimates more tentative. It seems that in general N_e is not much greater than 5×10^{10} cm^{-3} but definitely greater than 10^{10} cm^{-3}.

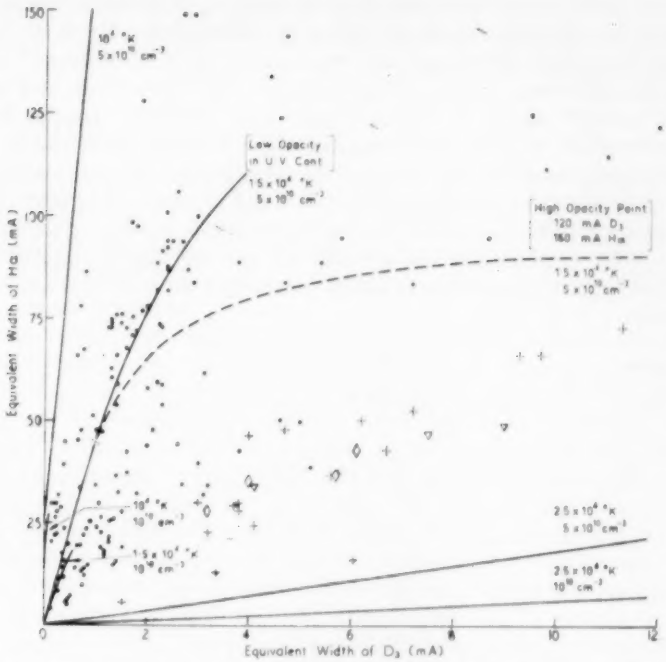


FIG. 1.—*Comparison of computed equivalent widths of Hα and D_3 with observations of Brück and Moss on a number of prominences. Points for some particular prominences are indicated separately. The broken part of the curve corresponds to the region of transition between high and low opacities in the U.V. continuum of the He I.*

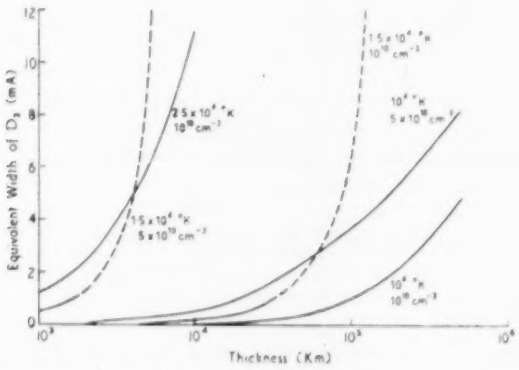


FIG. 2.—*Computed curves of D_3 equivalent width against model thickness. The broken parts of the curves refer to the region of transition between high and low opacities in the U.V. continuum of He I.*

5. *Temperature variations in a single prominence.*—A third set of observations by Brück and Moss (7) gives Hα and D_3 equivalent widths for a large number of points in a single prominence. These results are shown in Fig. 3 together with the computed curves extracted from Fig. 1. Points relating to different

regions are differentiated as explained in an insert and the Hα isophotes of the prominence, as given by Brück and Moss (**7**) are also shown. It is clear that there exist systematic differences and from the trend of the curves it could be inferred that the left-hand edge of the prominence has a rather higher temperature than the central regions. The alternative of a constant temperature and significantly lower electron concentration seems unlikely as it would require too great a thickness at the edge. On the right-hand edge, however, the behaviour is not as clear; the lower intensities of both Hα and D_3 point to a smaller thickness or electron concentration, or both, but there is no clear indication of a temperature variation. The separation of the prominence into two distinct regions may be seen from Fig. 3 by comparison of points for the left- and right-hand sides of the prominence.

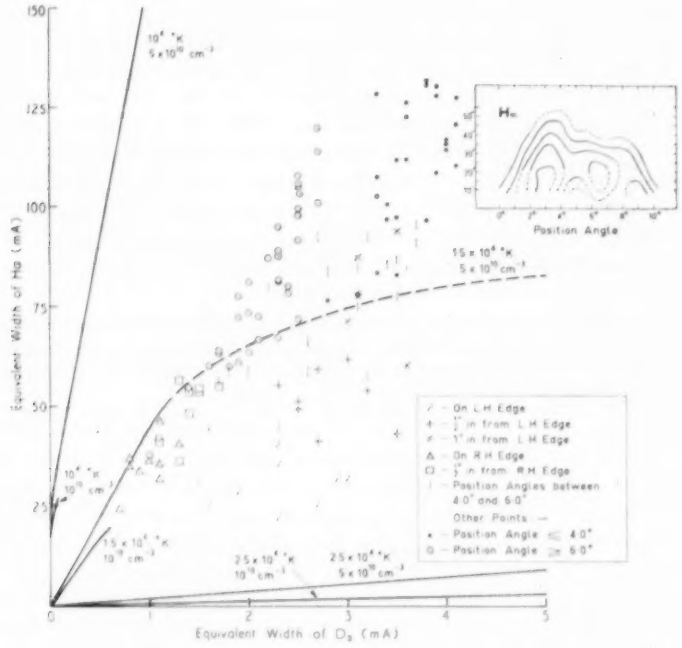


FIG. 3.—*Comparison of computed equivalent widths of Hα and D_3 with observations of Brück and Moss on a single prominence. Points in different regions of the prominence are differentiated as shown in the lower insert. The Hα isophotes of the prominence, as given by Brück and Moss* (**7**) *are shown in the upper insert.*

6. *Discussion.*—The conclusions reached here concerning the electron concentration and temperatures of an average prominence agree well with those found from previous analysis (**1**) of the Hα emission line contours and so suggest that the adopted helium abundance is probably not far out. In fact, by combination with the results found in (**1**), one may make an estimate of the relative helium abundance as follows.

For partially coherent scattering, it appears from (**1**) that the temperature range for all prominences is from about $1{\cdot}0 \times 10^4$ deg. K to $2{\cdot}0 \times 10^4$ deg. K.

Since a change in the relative helium abundance would change the D_3 equivalent width by an almost equal factor while leaving the Hα intensity unchanged we see from Fig. 1 that, for $T = 2 \times 10^4$ deg. K, a reduction of the helium abundance by a factor of more than about 1·5 would certainly result in a computed curve lying above many of the points in Fig. 1. At the other end of the temperature range, an increase of the relative abundance by a factor greater than two would shift the curves for $T = 10^4$ deg. K, $N_e = 5 \times 10^{10}$ cm^{-3} to a position such that some observational points lay to its left. Since the curves in Fig. 1 are based on an abundance ratio He/H = 0·2, the comparison of the present results with those found in (**1**) indicates that the helium abundance in prominences relative to that of hydrogen lies between the limits 0·4 and 0·12, and these limits have been generously estimated.

For the higher temperature models discussed here, certain cases seem to be excluded by observation. As may be seen from Fig. 2, a model with $T = 1{\cdot}5 \times 10^4$ deg. K, $N_e = 5 \times 10^{10}$ cm^{-3} gives D_3 emission of the right order of magnitude for thickness $\lesssim 4 \times 10^3$ km. However, as the thickness increases, the opacity in the ultra-violet continuum of helium becomes of the order of unity and the computed equivalent width of D_3 increases sharply to values $> 10^2$ mA, whereas the maximum value observed by Brück and Moss is 16 mA. There is evidently some fundamental reason for the non-appearance of the high D_3 emissions, e.g. that higher temperatures are associated with lower values of N_e. It is of interest to note that a similar relationship was suggested from the analysis of (**1**). As mentioned above, the observations shown in Fig. 3 seem to indicate a temperature decrease away from the left-hand edge of the prominence. The intensities of Hα and D_3 also indicate that the thickness or electron concentration, or both, increase away from this edge. This is at least not incompatible with the suggested variation of temperature with electron concentration.

7. *Conclusions.*—By comparison of the observed and computed equivalent widths of the Hα and D_3 emission of solar prominences it has been shown that a consistent interpretation of the observations is possible in terms of a slab model prominence whose temperature is between about $1{\cdot}1 \times 10^4$ deg. K to $2{\cdot}0 \times 10^4$ deg. K and electron concentration—and thus hydrogen abundance—between 10^{10} cm^{-3} to 5×10^{10} cm^{-3} if the helium abundance is one-fifth that of hydrogen. These results are in good agreement with those obtained earlier from an analysis of the contours of the Hα line emitted by prominences. When taken together, these two analyses indicate that the ratio of the helium to the hydrogen abundance in a solar prominence lies in the range 0·12 to 0·4.

It has been found that the D_3 prominence emission is made up almost entirely of diffusely scattered disk radiation.

8. *Acknowledgment.*—The author wishes to thank Dr R. G. Giovanelli for his continued interest and advice.

Division of Physics,
Commonwealth Scientific and Industrial Research Organization,
Sydney:
1956 *April* 23.

References

(1) J. T. Jefferies, *M.N.*, **115**, 617, 1955.
(2) M. T. Conway, *M.N.*, **112**, 55, 1952.
(3) M. A. Ellison, *Pub. R. Obs. Edinburgh*, **1**, No. 5, 1952.
(4) J. T. Jefferies, *Aust. J. Phys.*, **8**, 335, 1955.
(5) H. A. Brück and W. Moss, *M.N.*, **103**, 258, 1943.
(6) H. A. Brück and W. Moss, *M.N.*, **105**, 17, 1945.
(7) H. A. Brück and W. Moss, *M.N.*, **105**, 282, 1945.
(8) R. v. d. R. Woolley and D. W. N. Stibbs, *The Outer Layers of a Star*, p. 172, Oxford, 1953.
(9) R. G. Giovanelli and J. T. Jefferies, *Aust. J. Phys.*, **7**, 570, 1954.
(10) G. Munch, *Ap. J.*, **102**, 386, 1945.

AN EXAMINATION OF THE OBSERVATIONAL EVIDENCE FOR THE ACCRETION THEORY OF THE SOLAR CORONA

D. E. Blackwell and D. W. Dewhirst

(Received 1956 September 19)

Summary

The predictions and requirements of the accretion theory of the solar corona are compared with the results of observation, with special reference to the electron distribution in the neighbourhood of the Sun and convection in the inner corona. It is concluded that, even with reasonable modification, the theory is unable to satisfy the requirements of observation, and that accretion is therefore not likely to be an important coronal mechanism.

---

1. *Introduction.*—The theory of the accretion of interstellar gas by stars has been developed in papers by Bondi, Hoyle and Lyttleton (**1, 2, 3, 4**), and these authors have applied the theory to the solar corona (**5, 6**). McCrea has given a valuable survey of the status of the general accretion theory as it appeared in 1953 (**7**). It is likely that the accretion process is an important one in astrophysics generally, but its importance in solar physics has never yet been adequately discussed. Criticism of the accretion theory and of its application to the Sun has been made, on theoretical grounds only, by Atkinson (**8**), Öpik (**9**), Pikelner (**10**) and Gething (**11**), and there has been further discussion of some general points by Hoyle and Lyttleton (**12**), and by others at the conference on cosmic clouds held in Cambridge in 1953 (**7**). In a later publication Hoyle (**13**) has developed the accretion theory and presented more observational material to support it, without altering the essential form of the theory as it was proposed in 1947 (**5**). However, in this paper we shall attempt to assess the value of this theory of the corona by a more rigorous comparison of its predictions with the results of observation, using new data concerning the electron density in the neighbourhood of the Sun which have recently become available. Except for some criticisms by van de Hulst (**14**), we feel that comparison with observation has hitherto been inadequate and it now seems opportune to review the situation. In making a comparison it is necessary to remember the basic assumptions in the theory, and to make a final assessment we need to consider all reasonable modifications of it.

We are here concerned primarily with investigating the plausibility of the accretion mechanism as a solution to the problems of the corona. In addition, however, we may think of the solar neighbourhood as the most suitable testing ground for the accretion process in general astrophysics.

The relation of the observations to other theories of the corona has not been examined by us in this paper, and we do not wish to imply, by this omission, that they do not have difficulties.

2. *The accretion theory.*—Broadly speaking, the solar corona can be described as a hot and optically thin atmosphere, not in radiative equilibrium with the photosphere: its temperature is about 10^6 degrees. While there are features

of coronal structure which are still far from being understood, e.g. the fine structure of the inner corona, the chief problems are to understand the mechanism which maintains the high temperature of the inner parts, the reason for the great extension along the ecliptic, and the forces which so closely link the shape of the corona with the phase in the sunspot cycle. The accretion theory supposes that the energy for the maintenance of the coronal temperature comes from the gravitational energy of accreted gas, and that this gas is observed in the outer parts of the corona. The following is a summary of the theory based chiefly on references (**5, 6, 13**).

The Sun is supposed to be moving with velocity v through an interstellar gas cloud, which is chiefly hydrogen, containing τ_∞ atoms cm^{-3}. What happens when the Sun first enters this cloud is not clear, but it is supposed that after a short time an equilibrium state is reached. A tail is formed behind the Sun which traps hydrogen atoms or ions passing through it by reducing their transverse velocities. If the remaining radial velocity of an atom, after collision at a certain position along the tail, is less than the escape velocity from the Sun at that point, the atom will return to the Sun along an undetermined path and be captured.

This is a mechanism for converting a hyperbolic orbit into a parabolic or elliptical orbit. Three simplifying assumptions are now made. First, the gas pressure is never so great that pressure effects have to be taken into account; second, the flow of accreted atoms into the Sun is radial for all distances from the Sun of less than 1 A.U.; third, the effects of magnetic fields of solar and interplanetary origin are neglected. If M is the mass of the Sun and G the constant of gravitation, and all the interstellar gas within a distance σ of the Sun is captured,

$$\sigma = \frac{2MG}{v^2}. \tag{1}$$

Further, the rate of accretion of hydrogen atoms by the Sun is

$$\frac{dN}{dt} = \frac{4\pi G^2 M^2}{v^3}\tau_\infty. \tag{2}$$

When the Sun first enters the cloud the accreted material increases the temperature of the Sun's outer parts to form a more extended atmosphere. Further atoms on arrival are increasingly retarded by collision as they pass down through this atmosphere, until an equilibrium state is reached in which, at a distance r^* from the centre of the Sun, their momentum has been reduced to a fraction $1/e$ of its original value. For convenience, those atoms which still retain most of their momentum are called "accreted" atoms, whilst those which have undergone numerous collisions in this inner zone are called "solar" atoms. By making assumptions about the rate of loss of energy by impact, r^* is calculated to be $1{\cdot}17R_\odot$; that is, most of the energy of the incoming atoms is lost at a distance of about $0{\cdot}17R_\odot$ above the photosphere. It is shown that at this distance the density of incoming atoms is small in comparison with the density of atoms belonging to the solar atmosphere.

Besides providing energy, the incoming atoms transfer downward momentum to the solar atoms and in this way effectively increase the value of the gravitational acceleration. Let the density of accreted atoms at distance r^* be τ and let there be $N(r^*)$ solar atoms in a column of cross-section 1 cm^2 above this level. Then it is shown (**6**) that, because of the collisions they suffer with the accreted atoms,

the solar atoms are subjected to an effective gravitational acceleration $g(1+q)$, where g is the undisturbed gravitational acceleration due to the Sun, and

$$q = \frac{\tau V^2}{gN(r^*)}. \tag{3}$$

V is the velocity that an atom acquires at distance r^* in the gravitational field of the Sun

$$V = \sqrt{\left(\frac{2GM}{r^*}\right)}. \tag{4}$$

Much of the attraction of the theory arises from the numerical value of V, 570 km sec^{-1}. This allows ample margin to account for a supposed r.m.s. velocity of hydrogen atoms in the corona, namely 90 km sec^{-1} at 10^6 degrees.

From a comparison of the observed values of electron density in the chromosphere and corona with the computed values for various values of q, Hoyle deduces that $q = 0{\cdot}25$.* Hence he deduces τ, the density of the incoming atoms. Now, as it is assumed that all of the atoms contained within a cylinder of cross-section $\pi\sigma^2$ are swept up by the Sun and produce a density τ near the surface, we can calculate the density of hydrogen atoms at distance σ, considered to be the density at distances from the Sun large in comparison with the size of the solar system. Hoyle (**6**), assuming $v = 10$ km sec^{-1}, derives

$$\tau_\infty \sim 45 \text{ atoms cm}^{-3}$$

which is regarded as "well within the range given by studies of galactic structure and stellar evolution". In ref. (**13**) the value $\tau_\infty = 30$ atoms cm^{-3} is also given.

The kinetic energy of the incoming atoms is liberated by collision, chiefly at the height of $0{\cdot}17R_\odot$ above the photosphere, and it has now to be dissipated. According to Hoyle (**6**) dissipation by radiation or conduction is completely ineffective, and he therefore postulates a powerful system of convection currents operating between this layer and the photosphere. It is shown that the velocity of the downward current is of the order of 470 km sec^{-1} and that of the upward current is of the order of 750 km sec^{-1}, both velocities being relative to that of the incoming accreted material.

We notice at this stage that if this analysis is correct the total flux of energy into the corona from the accretion process is apparently much greater than is necessary to maintain the corona. The most reliable estimate of the rate of loss of energy from the corona is about $1{\cdot}4 \times 10^{27}$ ergs sec^{-1} (**15**). As 80 per cent of this energy loss is by conduction down to the photosphere, it is unlikely that this estimate will be greatly affected by uncertainties in the amount of radiation in the far ultra-violet. The total flux of energy from accretion is

$$\tfrac{1}{2}m_p V^2 \frac{dN}{dt} = \frac{4\pi G^2 M^2 m_p}{Rv^3}\tau_\infty = 0{\cdot}71 \times 10^{27}\tau_\infty \text{ ergs sec}^{-1},$$

where m_p is the mass of the proton, and v is assumed to be 10 km sec^{-1}. When τ_∞ is taken to be 45 atoms cm^{-3} this flux of energy is greater than the loss from the corona by a factor $\times 22$. The unwanted energy is convected down into the photosphere.

* Recent advances in knowledge of electron densities in the neighbourhood of the Sun require a modification of the value of q. We do not however calculate a new value because our later arguments are independent of a precise knowledge of this parameter.

In the following discussion we first compare (in Section 3) the requirements of the above theory (**5, 6**) with the results of observation, and then (in Section 4) see if this theory can be made more convincing by alteration of detail.

3. *Comparison with observation.*—In this first analysis we use the values of the parameters given by Hoyle (**6**) viz.

$$v = 10\,\mathrm{km\,sec^{-1}},$$

$$\tau_\infty = 44{\cdot}8\,\mathrm{atoms\,cm^{-3}}$$

giving

$$\sigma = 3{\cdot}82 \times 10^3 R_\odot = 17{\cdot}8\,\mathrm{A.U.}$$

The theory is concerned with numerical values of the density of interstellar hydrogen close to the Sun; observationally we can study only the electron density. To compare theory and observation we therefore need a knowledge of the degree of ionization of the interplanetary hydrogen as a function of distance from the Sun. There is no doubt that at distances of the order $10R_\odot$ the ionization is virtually complete. At distances of the order of $100R_\odot$ there is some uncertainty. Hoyle (**6**) has suggested a value of the order of 50 per cent; Öpik (**9**), using a value for the flux of ionizing radiation obtained from ionospheric studies, derives a value of 99·7 per cent at $200R_\odot$. In the following discussion we assume, with Öpik, that the ionization is complete in the region which concerns us, but if the degree of ionization is not less than 50 per cent at $100R_\odot$ from the Sun our arguments are little affected.

(*a*) *The electron density near the Sun.*—In the region of the corona where most of the accreted material is stopped by impact with solar material, one might expect a change in the gradient of electron density. Hoyle writes "Actually, such a region does exist, at about twice the photospheric radius from the centre of the Sun. The density of the material of the solar atmosphere falls off steeply at this distance" (Hoyle (**13**), p. 119). It seems to us that there is no observational support of any kind for this statement. It is true that the electron density varies with position angle in a rather irregular way, corresponding to the complex structure of streamers in the inner corona, but there is apparently no position within $10R_\odot$ of the Sun where the electron density changes in the way suggested. Fig. 1 shows the gradient of electron density derived from the analysis of van de Hulst (**16**) for the equatorial regions at sunspot maximum: at this time there is little difference between the equatorial and polar regions. The change of gradient with distance is regular and there is no special anomaly either at $1{\cdot}17R_\odot$ or at $2{\cdot}0R_\odot$. At sunspot minimum in the polar regions there is no evidence for an abrupt change of electron density out to the observable limit of $6R_\odot$: in the equatorial regions a marked change of gradient occurs only at about $25R_\odot$ from the Sun (**18**). It is clear that there is no support here for the accretion theory.

(*b*) *The electron density in the outer corona.*—(i) The assumption that the flow of accreted material into the Sun is isotropic for all distances up to 1 A.U. is evidently an oversimplification. At sunspot minimum the electron density in the polar regions is considerably smaller than that in the equatorial regions. At $6R_\odot$ from the Sun the electron densities are probably in the ratio of 1 : 25 (**17, 18**). Were the theory correct, most of the electrons at this distance would be accreted electrons, and it would seem reasonable to deduce that accretion in

the polar regions is negligible. But in spite of this the chromosphere still exists in the polar regions at sunspot minimum, and, for the extreme inner corona, the tables of van de Hulst (**16**) show only a small difference between the gradients of electron density at the pole and at the equator. The diagrams of the brightness of the 6374 A FeX emission line plotted by Waldmeier (**19**) show that there is often emission in this line at the poles near sunspot minimum: evidently the temperature of the corona here cannot be much less than 10^6 degrees. However,

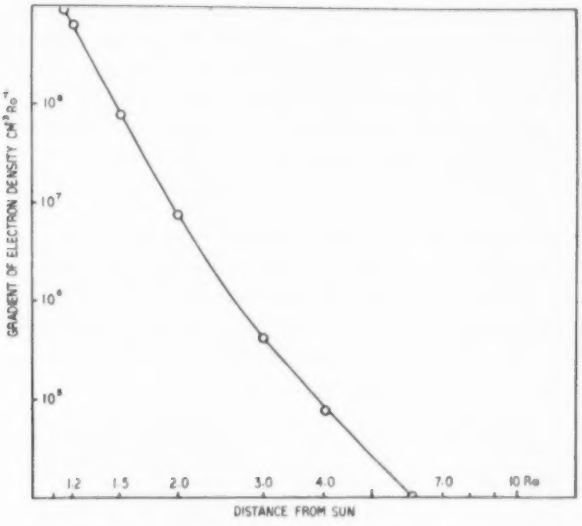


FIG. 1.—*Gradient of electron density in the solar corona. Data from van de Hulst, referring to the equator at sunspot maximum.*

it may be reasoned that even if there is no accretion at the poles a corona would still be expected to exist there because of the known high thermal conductivity of these regions. But we notice now that in the absence of a magnetic field one may expect the difference in electron density between pole and equator to be appreciably diminished by diffusion in a time $2R_\odot/u$, where u is the average thermal velocity of a proton in the corona: this time is of the order of 3 hours. In fact, the large difference in density appears to exist for at least 5 years during sunspot minimum. There is therefore some force preventing mass diffusion along the density gradient from the equator to the poles. If mass diffusion is inhibited then thermal diffusion must also be inhibited.

Accretion, if the dominating mechanism, must somehow be reduced near the poles. It is tempting to link this with the presence of a general magnetic field. There are three points here, however. First, a magnetic field of the dipole type might be expected to *increase* the rate of accretion of charged particles at the poles. Secondly, if a magnetic field is admitted to be responsible, an important factor has been omitted from the accretion theory. Thirdly, if, as the theory suggests, the electron density at $6R_\odot$ from the Sun is due solely to accretion why should this density change so markedly (over a range of 25 : 1) with the solar cycle?

(ii) The situation in the equatorial plane is no more promising. Near the Sun, in the region ordinarily observed at eclipses, the electron component of the corona is nearly symmetrical about the solar equator. But far from the Sun, in the zodiacal light region, it is more nearly symmetrical about the ecliptic. It has been shown (**20**) that between elongations 25° to 43°, corresponding to minimum distances from the Sun of $90R_\odot$ and $146R_\odot$, the angle between the axis of symmetry of the electron component and the ecliptic is probably less than 1°. At some distance from the Sun between $90R_\odot$ and $10R_\odot$ the axis of symmetry must change from the ecliptic to the solar equator, but the precise position is unknown. There is no correspondence at all between this observational picture and the picture presented by the accretion theory. The symmetry of the electron component about the ecliptic suggests that these electrons are at least semi-permanent members of the solar system. Even if the direction of motion of the Sun through the supposed gas cloud were in the plane of the Earth's orbit it would be difficult to see how the observed symmetry of electron distribution could be produced.

(iii) An important piece of evidence quoted as being in favour of the accretion theory is the variation of electron density with distance from the Sun. The original theory assumes free fall with no pressure effects and no collisions with the solar atmosphere beyond about $2R_\odot$ from the centre. In this case the electron density should vary with distance according to the law

$$N = kr^{-3/2}. \tag{5}$$

The density of the *accreted electrons* at $1{\cdot}2R_\odot$ has already been found by assuming $q = 0{\cdot}25$ (p. 639). This value may now be used to predict the electron density at any distance from the Sun. Originally, the electron densities predicted by the theory were compared with the Baumbach model of the corona. This much-quoted model, which took no account of the F (dust) corona, was essentially a summary of all measures available at the time of its construction twenty years ago; it must now be discarded in favour of more recent measures (**17, 18, 20, 21**) in whose reduction full allowance has been made for dust. Since observations show the K (electron) corona to be far from radially symmetrical, comparison with any theory assuming radial symmetry is of necessity somewhat artificial. In Fig. 2 we plot the observed electron densities near the ecliptic on the same diagram as the predictions of the accretion theory. Two predictions are given here. The upper line assumes with Hoyle (**6**) an electron density of $2{\cdot}1 \times 10^6$ cm^{-3} at $1{\cdot}17R_\odot$. Then at $20R_\odot$ the predicted electron density is too great by a factor of $\times 13$, although at $150R_\odot$ the discrepancy has been reduced to a factor of 2·5. If we fit the theory to the measures at $150R_\odot$ the overall picture is naturally improved (middle line) but the observed gradient of electron density at $150R_\odot$ is much less than is predicted. The observed electron densities of Fig. 2 have been calculated by making assumptions about the polarization of the F component. These assumptions have been varied as far as seems reasonable to give maximum and minimum values outside which the actual electron density is unlikely to fall (**20**), and these extreme values are plotted in Fig. 3 along with the two predicted curves already given in Fig. 2. Although predicted electron densities are nowhere more than three times greater than the observed values, there is a great difference between the observed and predicted gradients at $4R_\odot$ and $150R_\odot$.

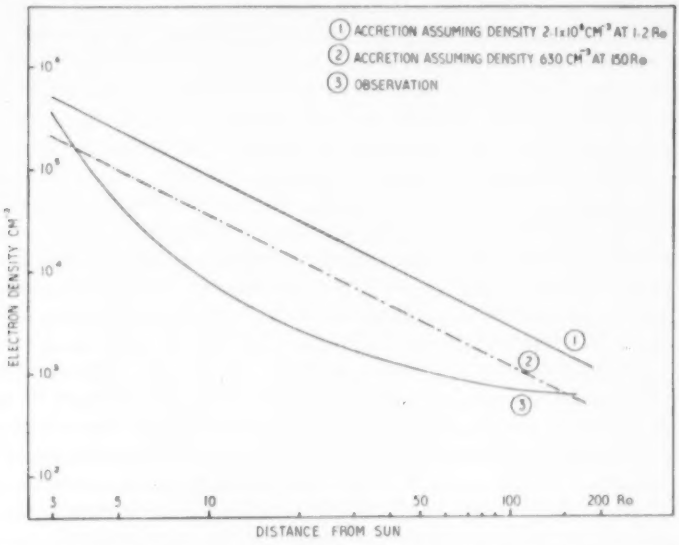


FIG. 2.—*Comparison of observed electron densities near the ecliptic with those predicted by the accretion theory.*

(*c*) *Convection in the inner corona.*—The original accretion theory supposes that the whole of the solar atmosphere between the chromosphere and the part of the corona where most collisions occur is in violent convective motion. This motion is necessary to remove accretion energy from those parts of the corona where it is liberated and is shared by all parts of the region. It is nearly radial and attains shearing velocities of about 280 km sec^{-1}.

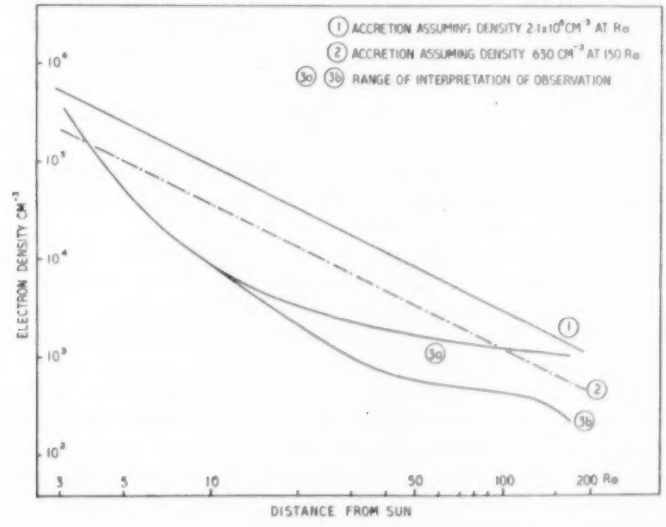


FIG. 3.—*Comparison of maximum and minimum likely electron densities near the ecliptic with those predicted by the accretion theory.*

Here there is great disagreement with observation. According to Kraul (**22**) the width at half intensity of the 5303 A line in the coronal spectrum at the height of 0'·5 above the limb ($\equiv 23\,000$ km or $1{\cdot}03R_\odot$) is about 0·7 A. The measurements of von Klüber (**23**) show that up to a height of about 4' the width is of the order of 1 A. Waldmeier (**24**) has considered the possibility that the contour of the lines is due to the streaming of coronal material away from the Sun, and has carried through the necessary integrations along the line of sight. An extension of his computations shows that the whole of the observed line width can be explained as due to ascending and descending columns in the region with speeds of 47 km sec^{-1} relative to the Sun. In fact, it is probably not permissible to ascribe more than 10 per cent of the line width to motions of this kind, the remainder being due to thermal motions corresponding to a temperature $\sim 2 \times 10^6$ degrees. The maximum permissible convection is therefore* of the order of ± 5 km sec^{-1}. Even remembering that Doppler broadening is not a very sensitive method of detecting convection currents normal to the photosphere, it is difficult to reconcile the observed widths of coronal lines with the powerful and slightly non-radial convection currents postulated by the theory. A more promising method of seeking evidence for convection normal to the photosphere is to look for proper motions of coronal material either in the 5303 A line using a coronagraph, or in white light during eclipses. Van de Hulst (**25**) has reviewed the published information about these motions. With a coronagraph it is possible to see fine structure in the corona up to heights of about $0{\cdot}3R_\odot$. Both Lyot and Waldmeier have reported coronal condensations which remain identifiable for periods of 6 hours or longer, during which time the features show small positional displacements with velocities only rarely in excess of 10 km sec^{-1}. In eclipse studies in white light fans and arches higher in the corona have been followed for periods of up to $2\frac{1}{2}$ hours, and have shown little motion or change in shape. On the whole such motion as exists tends to be away from the photosphere. It is difficult to see how such structures could retain their identity for such long periods were they subject to shearing velocities of 280 km sec^{-1}. Under these conditions a coronal condensation of diameter 10" would double its size in 30 seconds.

(*d*) *The hydrogen density in the neighbourhood of the solar system.*—In order to explain the order of magnitude of the observed electron densities in the corona and zodiacal light regions, even assuming that the capture process is perfectly efficient, the theory requires the density of interstellar hydrogen in the neighbourhood of the solar system to be not less than 30–45 atoms cm^{-3}. We may enquire whether this value is a reasonable one.

The most generally accepted mean density of hydrogen in the galactic plane, based on observations of H II regions and on dynamical grounds, is about 1 atom cm^{-3}. There is good evidence that the distribution is not uniform. There are indications that near the galactic plane the hydrogen is concentrated into clouds of density about 10 atoms cm^{-3} which occupy 7 per cent of interstellar space; between the clouds the density is not greater than 0·1 cm^{-3} (**26**). Both optical and 21 cm measures show that the Sun is on the inner edge of a spiral arm, and the latter measures suggest that the mean density within 100 parsecs of the Sun is not greater than 0·4 atoms cm^{-3} (**27**). At present there appears to be no method, either optical or radio, of deciding whether or not the concentration of material in a region within 5 parsecs of the Sun is about

* This point is further discussed by von Klüber (**23**).

40 atoms cm^{-3}. We note, however, that the accretion theory not only requires the Sun to be inside a cloud, but also requires this cloud to be of somewhat greater density than the average.

(*e*) *Qualitative aspects.*—The qualitative aspects must bear less weight as evidence. However, it seems difficult to reconcile the usually complex aspect of the inner corona, so remarkably dominated by the sunspot cycle, or the remarkable symmetry of the inner corona about the equator at sunspot minimum, with the accretion theory.

4. *Possible modifications of the theory.*—The foregoing comparison with observation shows that the theory is unacceptable in its present state. We have now to consider whether the theory can be modified or whether it has to be rejected altogether. We consider it doubtful whether any modification of the theory can explain the observed distribution of electrons around the Sun, but it is worthwhile considering whether alteration of the parameters can remove some of the other conflicts with observation. In this discussion we restrict ourselves to the type of theory which has been outlined in Section 2, namely that in which the accreted material is gas of interstellar origin whose infall also maintains the energy balance of the corona.

(*a*) *Alteration of* v *and* τ_∞.—We have seen that if $v \sim 10\,km\,sec^{-1}$ and $\tau_\infty \sim 45$ atoms cm^{-3} the rate of accretion of energy by the Sun is about twenty times greater than is necessary to balance the loss of energy by the corona. This difference leads to the necessity of postulating violent convection in the lower corona. However, even without the introduction of an assumed value of the energy loss from the corona, the theory still requires convective transport of energy from the collision region, at about $1R_\odot$ above the photosphere, down to the lowest parts of the corona. It is important to see whether the situation can be improved by alteration of v and τ_∞. If we suppose that accretion is a perfectly efficient process and wish to account for the electron density at any particular distance from the Sun, we notice that v and τ_∞ cannot be chosen independently. For, suppose that the electron density at distance r_0 is fixed by observational requirements at N_0. Then, assuming symmetrical accretion for all distances less than r_0, the flux of energy into the Sun is fixed at

$$\frac{dE}{dt} = 2\pi m_p N_0 r_0^{1/2} (2GM)^{3/2} \tag{6}$$

independently of the value of τ_∞ and v that may be chosen later. To obtain τ_∞ and v we may now equate the flux of energy from (6) to that obtained from the accretion theory,

$$\tfrac{1}{2} m_p V^2 \frac{dN}{dt}.$$

Using the expression already quoted for dN/dt (equation (2)) we find

$$\frac{\tau_\infty}{v^3} = \frac{N_0 \sqrt{(2GM/r_0)}}{4\pi G^2 M^2}. \tag{7}$$

Hence a choice of v automatically leads to τ_∞ being fixed, and vice versa.

Unfortunately it is not possible to measure the velocity, with respect to the Sun, of the gas in our immediate neighbourhood. Adams (**28**) has made many observations of the radial velocities of distant clouds by measurements of the Doppler shifts of interstellar lines. These observations have been rediscussed by Blaauw (**29**) who has shown that the velocities of the clouds fit a distribution

of the form $e^{-\eta/v}$, where η is the mean velocity, equal to 5 km sec^{-1}. As the solar motion with respect to nearby stars is of the order of 20 km sec^{-1} it seems most reasonable to take v as being of this order of magnitude rather than 10 km sec^{-1}. If we take v to be 20 km sec^{-1}, then τ_∞ must be 360 atoms cm^{-3}—which is certainly an unreasonably high value. Only in this way can the electron density at $150R_\odot$ be reconciled with the accretion theory. When $v = 10$ km sec^{-1} we have for the capture radius $\sigma = 17{\cdot}8$ A.U. and for $v = 20$ km sec^{-1}, σ is 4·5 A.U. If it is supposed that the region of symmetrical accretion is scaled down in proportion to the capture radius, then for the latter case it would be of radius $54R_\odot$, or 0·25 A.U. As observation can be carried to $150R_\odot$ we should expect to observe considerable asymmetry along the accretion axis in the zodiacal light region. But there is no sign of an accretion axis. It is equally clear from observation that accretion is not symmetrical within the sphere of radius $x = 54R_\odot$.

If τ_∞ is reduced to a local mean value of 0·5 atoms cm^{-3} from Hoyle's value of 44·8 atoms cm^{-3}, we find that v must be reduced to 2·2 km sec^{-1} if the theory is still to give the required electron density at $150R_\odot$. But it is doubtful whether Hoyle's original theory is still applicable at such low velocities because pressure effects become important, and we should instead consider Bondi's theory (**30**) of symmetrical accretion which takes these effects into account. However, the observed departure from spherical symmetry is so great that this theory can have little relevance in this case.

(*b*) *Alteration of collision cross-section in the solar atmosphere.*—It has been suggested by van de Hulst (**14**) that r^* should be increased from $1{\cdot}17R_\odot$ to $3R_\odot$. However, the above analysis (equations (6) and (7)) shows that although r^* is altered the flux of energy into the Sun and the choice of τ_∞ and v are unaffected by the particular value of the collision cross-section adopted.

(*c*) *Alteration of assumptions about symmetry.*—In the accretion theory as outlined in Section 2 the simplifying assumption of spherically symmetrical accretion was made to make the analysis more tractable. If any mechanism could be found which would force the spherically accreted atoms into the plane of the ecliptic as they approach the Sun, several of the difficulties we have referred to would be removed: the high electron densities in the plane of the ecliptic could be explained without postulating a high overall rate of accretion and a large density of hydrogen in interstellar space. It is not easy to conceive such a mechanism without postulating the existence of a general interplanetary magnetic field, for which there is no supporting evidence.

5. *Conclusion.*—We may summarize the reasoning as follows. Three aspects of the corona concern us: its high temperature, its extension along the ecliptic, and the variation of its shape during the sunspot cycle. An acceptable theory of the corona must provide a numerical value of the temperature, quantitative data on the change of electron density with increasing distance from the Sun, and a mechanism linking the distribution of electron density near the Sun with the cycle of solar activity. In 1947 no satisfactory explanation of all these features was available. The idea that there might exist an important interaction between stars and interstellar gas was then receiving much attention and it was natural to apply the concept of accretion to the coronal problem. At first sight the accretion theory of the corona is attractive because it explains at once both the high temperature and the outer extension of the corona, although it fails to explain the marked variation with sunspot cycle. Further examination, however,

shows that there are difficulties in accepting the theory as first proposed; difficulties which have been rendered more serious by recent observations. First, the theory does not describe, even qualitatively, the electron distribution within $150R_\odot$ of the Sun. Second, to account even approximately for the electron densities at large distances from the Sun, it is necessary to postulate a high rate of accretion which leads to the supposition of powerful convection currents in the inner corona. There is strong evidence against the existence of any large convective energy transfer in this region. Third, it is shown that the theory leads to a relation between τ_∞ and v which is independent of the nature of the interaction between the solar atmosphere and the accreted gas. It seems impossible to assign a reasonable value to one of these quantities without deriving an implausible value for the other.

We conclude that, whilst it is not impossible that some part of the coronal material might be of interstellar origin, it is unlikely that a theory of accretion can explain either the density of the Sun's outer atmosphere or the temperature of its inner parts.

Mr Hoyle read the manuscript of this paper and made many helpful criticisms. We are also grateful to Professor W. H. McCrea for a valuable discussion on some general aspects of the accretion mechanism, and to Professor R. O. Redman for his encouragement and helpful criticism.

The Observatories,
Madingley Road,
Cambridge:
1956 *September* 18.

References

(1) F. Hoyle and R. A. Lyttleton, *Proc. Camb. Phil. Soc.*, **35**, 405, 1939.
(2) F. Hoyle and R. A. Lyttleton, *Proc. Camb. Phil. Soc.*, **35**, 592, 1939.
(3) F. Hoyle and R. A. Lyttleton, *Proc. Camb. Phil. Soc.*, **36**, 325, 1940.
(4) H. Bondi and F. Hoyle, *M.N.*, **104**, 273, 1944.
(5) H. Bondi, F. Hoyle and R. A. Lyttleton, *M.N.*, **107**, 184, 1947.
(6) F. Hoyle, *Some Recent Researches in Solar Physics*, Cambridge, pp. 39–83, 1949.
(7) W. H. McCrea, E. Schatzman, P. Ledoux *et al.*, *Gas Dynamics of Cosmic Clouds* (I.A.U. Symposium No. 2), pp. 186–211, Amsterdam, 1955.
(8) R. d'E. Atkinson, *Proc. Camb. Phil. Soc.*, **36**, 314, 1940.
(9) E. J. Öpik, *Zeits. f. Astrophys.*, **35**, 43, 1954.
(10) S. B. Pikelner, *Usp. Astron. Nauk.*, **6**, 281, 1954 (summarised Astron. News Letter No. 81, p. 1, April 15, Lille, 1956).
(11) P. J. D. Gething, *M.N.*, **111**, 468, 1951.
(12) F. Hoyle and R. A. Lyttleton, *Proc. Camb. Phil. Soc.*, **36**, 323, 1940.
(13) F. Hoyle, *Frontiers in Astronomy*, London, pp. 117–126, 1955.
(14) H. C. van de Hulst, *The Solar System*, Vol. 1, ed. G. P. Kuiper, Chicago, p. 313, 1953.
(15) R. v. de R. Woolley and D. W. N. Stibbs, *The Outer Layers of a Star*, Oxford, p. 234, 1953.
(16) H. C. van de Hulst, *The Solar System*, Vol. I, ed. G. P. Kuiper, Chicago, p. 262, 1953.
(17) D. E. Blackwell, *M.N.*, **115**, 629, 1955.
(18) D. E. Blackwell, *M.N.*, **116**, 56, 1956.
(19) M. Waldmeier, *Die Sonnenkorona*, Vol. 1, Zurich, 1951.
(20) D. E. Blackwell, *M.N.*, **116**, 365, 1956.
(21) A. Behr and H. Siedentopf, *Zeits. f. Astrophys.*, **32**, 19, 1953.
(22) W. Kraul, *Zeits. f. Astrophys.*, **33**, 174, 1953.
(23) H. von Kluber, *M.N.*, **115**, 343, 1955.

(24) M. Waldmeier, *Zeits. f. Astrophys.*, **15**, 44, 1938.
(25) H. C. van de Hulst, *The Solar System*, Vol. I, ed. G. P. Kuiper, Chicago, p. 293, 1953.
(26) C. W. Allen, *Astrophysical Quantities*, London, p. 225, 1955.
(27) H. C. van de Hulst, C. A. Müller and J. H. Oort, *B.A.N.*, **12**, 117, 1954.
(28) W. S. Adams, *Ap. J.*, **109**, 354, 1949.
(29) A. Blaauw, *B.A.N.*, **11**, 459, 1952.
(30) H. Bondi, *M.N.*, **112**, 195, 1952.

A CATALOGUE OF STELLAR-LIKE EMISSION OBJECTS IN THE SMALL MAGELLANIC CLOUD

E. M. Lindsay

(Received 1956 July 12)

Summary

The Catalogue gives the X and Y co-ordinates, photographic magnitudes and blue-red colour indices of ninety-seven emission objects found with the ADH telescope and objective prism in the Small Magellanic Cloud. An attempt has been made to classify the objects. It is suggested that part of the Cloud in the direction of the extension has the characteristics of a Type I population.

Introduction.—In a previous paper (**1**), to be referred to as Paper I, it was shown that thirty-three objects in the Small Magellanic Cloud having the N_1, N_2 nebular lines in their spectra can be divided into two groups. Thirteen objects in Group I are associated with nebulosity and are probably early type stars exciting this nebulosity. Seventeen objects in Group II have the characteristics of planetary nebulae. Three objects were too bright to measure on the then available plates. An additional eleven objects bright enough to show Hα emission only and otherwise having the same characteristics as Group II were assigned as possible planetary nebulae.

Catalogue.—The present paper contains a Catalogue of all stellar-like objects in the Small Magellanic Cloud found by the writer which show emission. A total of ninety-seven has so far been found with the ADH telescope and objective prism. The columns in the Catalogue give the Catalogue number, the positions in X and Y coordinates as described by Miss Leavitt (**10**), the blue photographic magnitudes, red colour indices (blue–red photographic), blue photographic absolute magnitudes and remarks which include the assigned classification.

Magnitudes.—The magnitudes are based on photoelectric measurements made by King (**2**) on the stars of Sequence I (**3**) of the small Cloud. Photoelectric magnitudes have also been obtained for most of these stars by Gascoigne and Kron (**4**). Comparison of the systems is given in Table I. The two systems are in best agreement at star q; systematic differences are found between the

TABLE I

Comparison between Harvard photographic magnitudes (H) and photoelectric magnitudes of King (K) and Gascoigne and Kron (GK) in Sequence I of the small Cloud

Stars	$H-K$	$H-GK$
k–p	−0·17	−0·25
q	+0·02	−0·04
<q	+0·20	−0·01

(King did not measure star k. Of the stars fainter than q, King measured r, t, v, w, x; Gascoigne and Kron measured r-v. Star v is not included in the Table. It is suspected by Gascoigne and Kron to be variable. On the Armagh plates it seems to be variable.)

magnitudes brighter and fainter than this. There is no way at present of choosing between them. In Paper I King's magnitudes were used with a correction of $+0^{m}\cdot2$ to the Harvard magnitudes fainter than star x. In the present paper the same magnitudes are used without a correction to the Harvard stars fainter than x. This gives a smoother calibration curve and is probably more in agreement with the system of Gascoigne and Kron. The red magnitudes are those determined by King and reduced by him to the international system.

All magnitudes had been estimated through a binocular microscope using a scale of graduated stellar images which was calibrated from the sequence stars for each plate. An Iris Diaphragm Photometer (**11**) has since been acquired at Armagh and magnitudes have been re-determined with this. In addition to being an objective method, the relation between diaphragm reading and magnitude is linear over a large magnitude range. On a 60-minute exposure, for example, it is possible to measure the magnitudes of emission objects over the range $m = 10$ to $m = 17$. In the mean the agreement between photometer measures and eye estimates is good. For blue magnitudes the difference is photometer − eye $= -0^{m}\cdot14$; for red magnitudes, photometer − eye $= -0^{m}\cdot08$. The differences depend on the magnitudes as shown in Table II.

TABLE II

Differences between magnitudes measured with a Haffner photometer and estimated by eye ($\Delta m = m_p - m_e$)

Range	11–12	12–13	13–14	14–15	15 16	16–17	17
Δm_{pg}	−0·46 (1)	−0·22 (8)	−0·20 (19)	−0·14 (22)	−0·06 (14)	−0·06 (9)	−0·01 (1)
Δm_{pr}	−0·39 (1)	−0·32 (7)	−0·07 (26)	−0·06 (19)	−0·02 (18)	−0·04 (2)	...

The table shows the well-known effect that the eye is better at estimating faint objects than bright ones. The same is true in making star-counts to different limiting magnitudes using a scale of images; the most consistent results, and of a high order of accuracy, are those down to the faintest limiting magnitude. Part of the differences in the brightest stars in the table is probably due to the photometer. With long-exposure plates on a dense field the brighter images cover a background of faint stars. The photometer will combine the total density giving a brighter magnitude. The magnitudes for stars involved in nebulosity (NGC 346, 456, 460) are not included in the table. For some plates the exposures were not short enough to eliminate the nebulosity and the photometer measures are too bright. In such cases eye estimates were adopted.

The Iris Diaphragm Photometer operates in an unambiguous way in the photometry of point-sources whose photographic images are built up by diffusion of light in the photographic film and optical aberrations. When the source is not point-like, e.g. in the case of a star involved in nebulosity, the reading of the iris diaphragm cannot be interpreted in simple terms. Eye estimates depend directly on diameter, and, provided the nebulosity is not too dense, there is no difficulty in separating the star image from the background.

The correction for distance from the centre of the plate has been re-determined with the photometer using a series of exposures of HSR E6 at different distances from the centre. No correction was necessary within a radius of 7·5 cm from the centre. Most of the emission objects are contained within this and the necessary small corrections were applied to the few outside it.

The material for determining the magnitudes listed in Paper I was increased by an additional twenty 103a–O plates, ten of 5-min exposure and ten of 10-min exposure. These were taken for a joint programme with Dr Shapley for establishing further magnitude sequences in the Small Cloud. From the results of the different plates it is found that the probable error of a single plate is $\pm 0^m{\cdot}09$ for blue plates and $\pm 0^m{\cdot}15$ for red plates. The larger error for the red plates is chiefly due to the larger scatter of the sequence stars about the calibration curve. With the large number of plates used the probable error of the mean is smaller than that of the magnitudes of the sequence stars.

Catalogue

No.	X	Y	m_{pg}	C.I.	M_{pg}	Remarks
1	5100	11700	...	...	...	This object is in the direction of (or in) 47 Tucanae. The region is too crowded to be sure of the identification of the star and no magnitude measurements were possible. Position approximate. Lines Hβ, Hγ, Hδ. Type Be
2	6552	13260	16·39	+0·49	−2·3	Paper I. Planetary
3	9264	9642	16·35	+0·48	2·4	Paper I. Planetary
4	9522	5916	16·46	+0·84	2·2	Paper I. Planetary
5	9792	10092	16·38	+0·54	2·3	Paper I. Planetary
6	9918	8652	14·03	+1·20	4·7	Paper I. Planetary
7	9948	8448	17·4	+1·3	1·3	The faintest emission object discovered. Hα appears to be variable. Found on two plates only. On one it is noticeably bright and on the other it is barely visible and much fainter than No. 8. Long period variable or planetary ?
8	10008	7200	14·25	−0·11	4·3	This is in No. 242 of the Shapley–Wilson Catalogue (**8**) and is described there as " two nebulous stars ". It was considered by them to form with No. 241 part of the same nebula. Nos. 241 and 242 are Nos. 13 and 14 in the Catalogue of Virginia Nail, Whitney and Wade (**9**). The former is described as " several stars involved in nebulosity "; the latter as " nebulosity between two stars ". On the Armagh plates these two objects appear to be a small open cluster. The appearance of nebulosity seems likely to be due to faint stars. No. 8 is taken to be a Be star in the cluster. Hα emission weak
9	10242	9894	13·52	−0·34	5·2	Hα emission moderately faint. Probably Be
10	10284	8190	15·19	+0·58	3·5	Paper I. Probably a planetary
11	10338	7800	14·36	+0·38	4·3	Paper I. Planetary
12	10440	8520	13·75	−0·56	5·0	Paper I. Early type in nebulosity
13	10464	7428	13·52	−0·47	5·2	Paper I. Early type in nebulosity
14	10704	8454	12·96	−0·39	5·7	Paper I. Early type in nebulosity
15	10800	8328	12·53	+0·80	6·2	There is a line to the violet of Hγ which is probably λ4234 of Fe^+. If so, the lines are Hβ, Hγ, λ4234, Hδ and possibly Hε. Hα is very strong in this object. It is apparently a Be star, the red colour index being due to the great strength of Hα

CATALOGUE—(*cont.*)

No.	X	Y	m_{pg}	C.I.	M_{pg}	Remarks
16	10872	7980	13·26	+0·85	−5·4	Although the continuum is moderately strong no absorption is detectable. It does not resemble a late type. Hα emission is moderate and does not seem sufficient to account for the large positive red colour index. There is no other emission. Unclassified
17	11142	7980	13·51	−0·82	5·2	An H II sphere. Only Hα emission
18	11142	7944	14·80	+1·14	3·9	Paper I. Planetary
19	11142	9828	13·45	−0·18	5·2	Hα emission moderately weak. Probably Be
20	11244	8946	16·28	+1·01	2·4	Paper I. Probably planetary
21	11310	7446	15·46	+0·37	3·2	Paper I, where it is listed as a possible planetary. The magnitudes are uncertain since it is blended with a star. Measurements with the Haffner photometer result in a negative colour index. It may be a Be star in a cluster. The magnitudes and C.I. given here are from eye estimates. Hα emission is somewhat doubtful. Better to leave unclassified
22	11328	8394	13·71	−0·46	5·0	Emission somewhat doubtful but probably there. Apparently the central star of N 122 (HC 276) which is described as a " large faint nebulous patch, star involved " and N 34 (HR 377) denoted as an " arc of circle, central star ". This arc is the denser portion of more extended nebulosity. Probably Be
23	11340	8294	15·25	+0·34	3·4	It would seem that this should be included as a planetary. The lines N_1+N_2, Hβ, Hγ appear to be faintly present. Possibly λ3727. Continuum very faint
24	11490	7242	14·71	+1·66	4·0	Paper I. Planetary
25	11502	9534	14·97	+0·42	3·7	This seems to be another planetary. Hβ, Hγ, Hδ, λ 3727 faintly present with λ 3727 the strongest feature. Possibly N_1+N_2. Hα is roundish and diffuse. On most plates unseparated from two stars. Continuum very faint. The object is N 41 in H.R. 377 described as " small group nebulous (?) stars " and N 127 in H.C. 276 there described as " nebulous star or group of stars "
26	11568	7584	15·09	+0·80	3·6	Paper I. Probably planetary
27	11640	6180	16·01	+0·75	2·7	Paper I. Planetary
28	11718	8904	14·11	−0·76	4·6	Hα on only one plate and very weak. Probably Be
29	11736	9288	11·87	−0·60	6·8	Hα emission weak. Continuum fairly strong, partly overlapping another star with no sign of absorption. Probably Be
30	11838	7236	12·60	−0·53	6·1	Hα very weak if there at all. One broad emission line probably λ 4686. Type O
31	11850	9360	12·92	−0·15	5·8	Hα emission moderately faint. Probably Be
32	11946	10908	16·32	+0·71	2·4	Paper I. Probably planetary
33	11976	5388	15·96	+1·09	2·7	Paper I. Planetary
34	12012	14628	15·98	+0·69	2·7	Paper I. Planetary

CATALOGUE—(*cont.*)

No.	X	Y	m_{pg}	C.I.	M_{pg}	Remarks
35	12072	8040	14·24	−0·48	−4·5	Paper I. Early type in nebulosity
36	12072	9486	14·90	+0·75	3·8	Paper I. Planetary
37	12138	7614	16·07	+0·97	2·6	Paper I. Probably planetary
38	12438	8388	13·16	−0·14	5·5	Hα emission moderate. Probably Be
39	12558	10104	14·63	−0·38	4·1	Probably an H II sphere. Same as N 75 in H.R. 377, there described as a " round nebula, central star " and N 146 (H.C. 276) referred to as " nebulous stars ". It is possibly however an emission object in a small cluster of faint unresolved stars
40	12576	10128	14·61	−0·40	4·1	Close and similar to No. 39. May be in a faint cluster of unresolved stars. Probably an H II sphere. It is N 76 (H.R. 377) described as " a nebulous star " and N 147 (H.C. 276) described as " nebulous stars "
41	12696	9978	13·31	+1·03	5·4	Paper I. Probably planetary
42	12726	10740	14·73	+0·22	4·0	Paper I. Probably planetary. Blended with a star; magnitudes difficult to estimate
43	12900	9912	15·95	−0·45	2·8	There appears to be faint Hα emission on one plate. There is much overlapping of spectra in this condensed region. Probably Be
44	12990	12306	15·71	+1·28	3·0	Paper I. Probably planetary
45	13152	11964	13·06	−0·09	5·6	Early type from spectrum. Probably Be
46	13350	8748	13·25	−0·23	5·4	Hα emission fairly faint. Probably Be
47	13572	12408	13·00	−0·13	5·7	Early type from spectrum. Probably Be
48	13596	10548	14·08	+0·13	4·6	Paper I. Planetary
49	13788	9696	14·13	+0·84	4·6	Faint Hα emission appears to be present on two plates. Unclassified
50	13806	10158	14·64	−0·77	4·1	Paper I. H II sphere
51	13848	10098	14·34	+0·84	4·4	Paper I. Planetary
52	13926	10848	15·71	+1·43	3·0	Paper I. Planetary
53	14046	11868	12·15	−0·83	6·6	In NGC 346. Confused region of overlapping spectra and nebulosity. Magnitudes uncertain. Probably Be
54	14052	11856	14·71	−0·49	4·0	Same remarks as for No. 53
55	14070	11910	13·90	−0·23	4·7	Paper I. Early type in nebulosity (NGC 346)
56	14094	11772	16·07	+0·66	2·6	Paper I. Listed there as a possible planetary. Appears to be faint $N_1 + N_2$, Hβ on two plates. Planetary
57	14106	11838	15·04	−0·36	3·4	Paper I. Early type in nebulosity
58	14106	11922	14·63	+0·15	3·6	In NGC 346. It is doubtful what this object is. It was difficult to estimate magnitudes, especially on red plates on some of which it resembles nebulosity more than a star
59	14148	11928	11·20	−0·88	7·5	In NGC 346. Hβ, λ 4686. The latter very strong. A faint emission line to the violet of λ 4686 which may be λ 4542 of He^+. Listed by Miss Cannon (H.B. 897)
60	14160	12432	12·25	−0·20	6·4	Hβ, Hγ, Hδ. Hβ markedly strong. Type Be
61	14256	9162	14·00	−0·13	4·7	Hα emission very weak and on one plate only. Probably Be
62	14610	13980	15·08	+1·31	3·6	Paper I. Planetary
63	14958	11472	12·58	−0·05	6·1	Very weak Hα emission on one plate. Hγ, Hδ in absorption. Be

CATALOGUE—(*cont.*)

No.	*X*	*Y*	m_{pg}	C.I.	M_{pg}	Remarks
64	14970	9408	13·64	−0·24	−5·1	Hα emission very faint. Probably Be
65	15108	12896	14·19	−0·76	4·2	Paper I. H II sphere
66	15320	11652	13·48	+0·15	5·2	Hα emission moderately weak; continuum moderately strong. Unclassified
67	15564	9648	12·88	−0·32	5·8	Faint Hα emission on one plate. Spectrum partly overlapping another star but seems early type. Probably Be
68	15606	12450	13·67	−0·09	5·0	Very faint Hα emission on one plate. Probably Be
69	15684	11086	13·02	−1·12	5·7	Same as No. 68
70	15708	11104	13·56	−0·22	5·1	Very faint Hα emission. Probably Be
71	15732	12522	15·25	−0·06	3·5	In NGC 395. Dense region with nebulosity. Magnitudes difficult to estimate. On some plates, especially on red, this object appears as nebulosity. Probably early type involved in nebulosity
72	15744	12462	15·01	+0·06	3·7	In NGC 395. Also nebulous on some plates; magnitudes uncertain. Probably early type in nebulosity
73	15750	12486	14·75	−0·15	4·0	Same as Nos. 71 and 72
74	15768	12558	14·39	−0·70	4·3	Same as No. 73
75	15948	13278	12·76	−0·06	5·9	λ 4686 seems present. There are apparently some faint emission lines but the spectra are rather broad and do not go deep enough for certainty. Hα emission is moderately strong. Type O
76	15990	8190	13·05	−0·39	5·6	Hα emission faint. Probably Be
77	15996	10870	13·43	−0·32	5·3	Same as No. 76
78	16086	11388	14·57	+0·29	4·1	Faint Hα emission. Continuum fairly strong. Unclassified
79	16650	8094	10·38	+0·88	8·3	Paper I. No eye estimates of this (I.C. 1644) were given there due to its brightness. It is difficult to classify it. Although listed in H.A. 60 as a gaseous nebula there is no sign of nebulosity. The magnitudes are internally very inconsistent
80	17304	5652	14·24	+0·54	4·6	Red magnitude from one plate only Unclassified
81	17832	7620	14·3	...	4·4	Paper I. Early type in nebulosity. Value of m_{pg} given here of little weight
82	17850	7764	13·35	+0·05	5·4	Nos. 81–89 inclusive are in NGC 456 and 460. The region is very nebulous. The magnitudes are highly uncertain owing to the presence of this nebulosity. They are no doubt high temperature stars exciting the surrounding nebulosity
83	17892	7674	14·17	−0·86	4·5	Paper I
84	17958	7752	15·16	−0·34	3·5	Paper I. H II sphere
85	18036	7578	11·60	−0·80	7·1	Paper I
86	18060	7452	14·17	−0·11	4·5	Looks like nebulosity on red plate
87	18054	7578	13·55	−0·45	5·2	Paper I
88	18072	7482	15·87	−0·11	2·8	Doubtful which of two objects is showing emission
89	18072	7494	12·75	−0·12	6·0	Paper I
90	18354	8082	15·83	+0·72	2·9	Paper I. Probably planetary
91	19740	7650	15·94	+0·73	2·8	Paper I. Planetary

CATALOGUE—(*cont.*)

No.	*X*	*Y*	m_{pg}	C.I.	M_{pg}	Remarks
92	20538	7896	12·22	+1·32	−6·5	Paper I. No magnitudes are given there. The object resembles No. 79. No evidence of nebulosity although listed as a gaseous nebula (H.C. 224). Hα extremely strong. Possibly in small cluster
93	22050	7128	13·25	+0·06	5·4	Hβ emission present faintly. Nos. 93–96 inclusive were not on the same plate as the red standard sequence. Red magnitudes for these stars were obtained from a single red plate using a transferred sequence
94	22122	9252	12·32	+0·07	6·4	Hα emission strong. Hβ faint. Be
95	22260	6720	12·12	−0·70	6·6	Listed by Miss Cannon (H.B. 891). Broad emission at λ 4686. Type O
96	22602	8820	12·23	−0·37	6·5	λ 4686. Type O. No red prism plates on this region
97	24660	1620	...	...	...	One broad emission line, probably λ 4686. Type O. No red prism plates and no chart plates for magnitudes

Discussion.—In Paper I the apparent modulus of the Cloud was taken to be 18·9 photographic. In the present paper $(m - M) = 18·7$ is used in accordance with the conclusions of de Vaucouleurs (**5**).

The majority of the hitherto unpublished objects in the Catalogue show Hα emission only with no details on blue or red (without filter) spectrum plates. These spectrum plates were obtained during the experimental stage with the prism when the emphasis was on the quality of spectra that could be obtained, and on methods of guiding. The spectra are much wider than is necessary with consequent bright limiting magnitude. Fainter spectra will be obtained when the opportunity arises.

Red colour indices of Hα emission objects of any class can have a wide range depending on the intensity of the emission; if the latter is strong enough such indices can be positive for early type stars. In assigning types where only absolute magnitude and colour index are available it has been considered that negative indices indicate B-type stars. In some cases with large positive colour indices it has not been possible to assign types. Table III summarizes the data of the Catalogue.

The brightest stars absolutely are type O. The Be stars and those in nebulosity exciting the nebular lines come next at a magnitude fainter. A further magnitude fainter are the stars in nebulosity showing only Hα emission and stars in H II spheres. The distribution with respect to absolute magnitude is given in Table IV. Nos. 79 and 92 have been included under N_1.

Of the eleven B stars probably in the Cloud and listed in the H.D. Catalogue, nine are on Armagh plates. Data on these are given in Table V. These values are given with some hesitancy since they are not considered to be of a high order of accuracy. They bear out, however, preliminary results quoted by Gascoigne (**4**) indicating the possibility of reddening of some of the B stars. There is no connection between colour index and position in the Cloud. The first five stars are in the southern portion of the main body, i.e. the bar, or core, H.D. 5030 and 5045 being close together. H.D. 6884 and 7099 are also neighbours, the

TABLE III

Emission Objects in the Small Magellanic Cloud

Type	M_{pg}	C.I.	No.
P	−3·3	−0·79	20
(P)	−3·3	−0·81	9
N_1	−5·1	−0·47	10
N	−4·2	−0·26	9
Be	−5·4	−0·19	25
O	−6·5	−0·51	5
H II	−4·2	−0·58	6
Uncl.	−4·0	−0·56	8

REMARKS.—P and (P) represent planetary and possibly planetary nebulae. N_1 represents early type stars in nebulosity showing nebular lines ; N, early type stars in nebulosity with Hα emission only. No. 81 is not included since no C.I. available. No. 1 not included in Be stars ; no magnitudes available. Of the H II spheres, three show nebular lines, three Hα emission only. Data on these refer of course to the central stars. Nos. 79 and 92 with $M_{pg} = -8{\cdot}3$ and $-6{\cdot}5$ and C.I. $= +0{\cdot}88$ and $+1{\cdot}32$ respectively are not included. It is too doubtful what they are. The mean absolute magnitude of all emission objects of Table III is $M_{pg} = -4{\cdot}4$.

TABLE IV

Distribution of Emission Objects according to M_{pg}

$-M_{pg}$	O	Be	N_1	N	H II	Uncl.	Σ	P(P)
<7·9	0	0	1	0	0	0	1	0
7·9–7·5	1	0	0	0	0	0	1	0
7·4–7·0	0	1	1	0	0	0	2	0
6·9–6·5	2	4	1	1	0	0	8	0
6·4–6·0	1	7	1	0	0	0	9	0
5·9–5·5	1	9	1	0	0	0	11	0
5·4–5·0	0	2	3	1	1	2	9	1
4·9–4·5	0	1	3	1	0	2	7	2
4·4–4·0	0	0	0	3	4	1	8	4
3·9–3·5	0	0	0	1	1	1	3	6
3·4–3·0	0	1	1	1	0	1	4	3
2·9–2·5	0	0	0	1	0	0	1	7
2·4–2·0	0	0	0	0	0	0	0	6
>2·0	0	0	0	0	0	1	1	0

TABLE V

Absorption B stars in the Small Cloud

H.D.	m_{pg}	C.I.	M_{pg}
4862	10·67	−0·33	−8·0
4976	10·87	−0·18	−7·8
5030	11·11	−0·08	−7·6
5045	10·81	−0·27	−7·9
5277	10·88	−0·17	−7·8
6884	10·64	−0·15	−7·9
7099	10·96	−0·23	−7·7
7222	10·82	−0·33	−7·9
7583	10·56	−0·32	−8·1

former having a positive colour index. H.D. 7583 is in NGC 456, a region marked by nebulosity. None of these B stars shows any emission.

The distribution of emission objects is shown in Fig. 1, which also shows the axial bar, asymmetrical inner arm and outer whorl as defined by de Vaucouleurs (6).

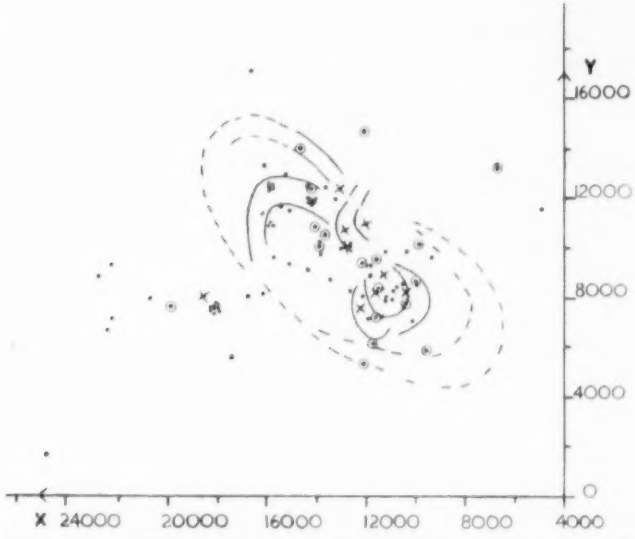


FIG. 1.—*Distribution of Emission Objects in the Small Magellanic Cloud with structural outline as given by de Vaucouleurs. Shapley's extension in the direction of the Large Cloud is towards the left.*

⊙ *Planetary* ;
× *Possible planetary.*

Of the sixty-eight non-planetary emission objects, nearly one-third are in clusters, NGC 346, 395 and 456 accounting for twenty. The majority of objects are contained in the bar, or core, of the Cloud. A high proportion of the planetary nebulae are however outside the core. Their distribution is shown in Fig. 2. Of the twenty planetaries, nine, or 45 per cent, are outside the region of greatest stellar density. Even if the probable planetaries are included, a considerable percentage is outside it. There is apparently an avoidance of the asymmetrical inner arm, as would be expected for members of Population II to which planetaries are known to belong.

The direction of Shapley's (7) extension towards the Large Cloud is of particular interest. Within this narrow belt are found twenty of the known emission objects. One of these is a planetary, one a possible planetary, leaving eighteen out of sixty-eight non-planetary in this comparatively star-poor region. Of this number there are three O-type stars, five early type exciting nebular lines and probably O-type, one H II sphere, seven showing Hα emission of which six are considered to be Be and one unclassified. The remaining two of the eighteen emission objects are Nos. 79 and 92 of the Catalogue, both of which are associated with nebular lines. Each appears to be in a small cluster but no nebulosity is visible. The absorption B star (H.D. 7583) is here. There are two regions of extended nebulosity. One is NGC 456. The other is at $X = 21720$,

$Y = 6300$ and seems to have remained undiscovered until now. This nebulosity, like that of NGC 456, appears to be at least as rich in blue as in red light. On a 90-min exposure 103a–E plate with red filter its largest dimension is over two minutes of arc but it is visible on a 10-min exposure 103a–O plate. The nebulosity in NGC 456 is stronger on a 30-min exposure 103a–O plate than on a 90-min exposure 103a–E plate with red filter. This part of the Cloud also contains fourteen open clusters, not counting the possibility of Catalogue Nos. 72 and 92 being each in one.

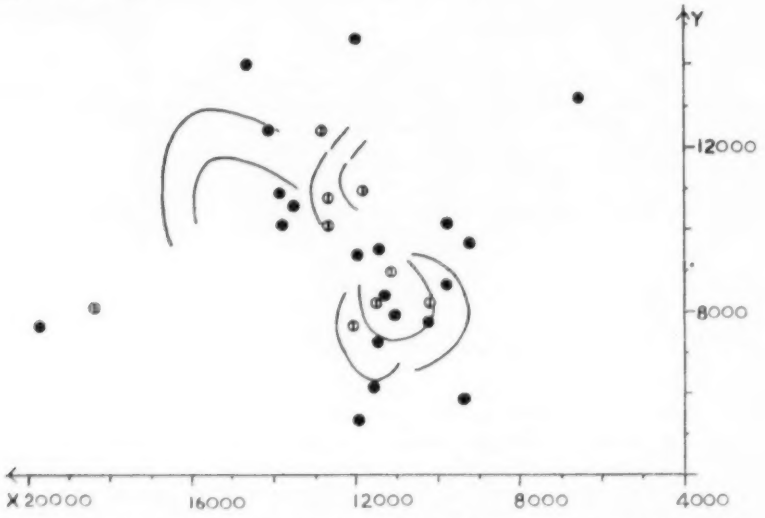


FIG. 2.—*Distribution of planetary nebulae* (●) *and possible planetary nebulae* (Ⓘ).

With a number of clusters, O, B stars and extended nebulosity it would appear that this part of the Cloud has many features in common with a Type I population. The present material does not cover the full region of the extension; the emission objects are within a diameter of just over six degrees. A programme will be carried through to include star-counts, colours and a spectral survey of the extension and a search for emission objects between the Large and the Small Clouds.

Armagh Observatory,
Armagh, N. Ireland:
1956 *July* 4.

References

(1) E. M. Lindsay, *M.N.*, **115**, 248, 1955; *Armagh Obs. Contr.* No. 17.
(2) Ivan King, *A.J.*, **56**, 43, 1951.
(3) Virginia McKibben, *Harv. Ann.*, **108**, 13, 1940.
(4) S. C. B. Gascoigne, *Suppl. Austral. J. Sc.*, **17**, No. 3, 1954.
(5) G. de Vaucouleurs, *P.A.S.P.*, **67**, 350, 1955.
(6) G. de Vaucouleurs, *Irish Astr. Journal*, **4**, 13, 1956; *A.J.*, **60**, 222, 1955.
(7) Harlow Shapley, *M.N.*, **94**, 791, 1934.
(8) Harlow Shapley and Harvia H. Wilson, *Harv. Circs.*, Nos. 275 and 276, 1925.
(9) Virginia McKibben Nail, Charles A. Whitney and Campbell M. Wade, *Proc. Nat. Acad. Sci.*, **39**, No. 12, 1953; *Harv. Repr.*, No. 377.
(10) H. S. Leavitt, *Harv. Ann.*, **60**, No. 4, 1907.
(11) H. Haffner, *Veröff. Sternw. Göttingen*, No. 106, 1953.

THE SENSE OF ROTATION OF NGC 253

David S. Evans

(Communicated by H. M. Astronomer at the Cape)

(Received 1956 October 23)

Summary

Low-dispersion spectra of five regions in the spiral nebula NGC 253 have been measured for radial velocity. The results indicate that the sense of rotation of the system is such that the spiral arms trail.

Photographs of the spiral system NGC 253, in three different ranges of wave-length, taken with the Radcliffe reflector, are reproduced in Plates 9(*a*, *b*) and 10(*c*). The northern and southern borders of this system differ markedly in appearance, the northern being characterized by the presence of a complex band of absorbing clouds. It is presumed that this edge is the nearer, so that the absorbing material is seen projected against the nebula.

The system is of low surface brightness, and early efforts, extending over some years, to obtain spectra, proved unsuccessful. In 1956 a single-prism Newtonian spectrograph designed by Dr A. D. Thackeray was constructed in the Royal Observatory workshops. Messrs Cox, Hargreaves and Thomson Ltd, supplied a collimator mirror, and were also the makers of the camera, originally used in a two-prism Newtonian instrument. This camera is of Schmidt type with a focal ratio *f*/1, and the spectra, with a dispersion of 400 A/mm at Hγ, are photographed on disks of film 8 mm in diameter. A first set of spectra was obtained on panchromatic film with the idea in mind of employing the emission lines thought likely to be present from a study of the photographs. Emission, showing a complex structure, was found near Hα. Difficulties were also encountered in connection with exposure times and standard wave-lengths for this region of the comparison spectrum. The situation was further confused by the presence of emission lines of terrestrial origin, including mercury and, possibly, neon from city lights and auroral lines. This set of films has, therefore, not been used.

A new set of films, using 103a–O emulsion, and with the mirror freshly silvered, was then obtained. The comparison spectrum was provided by a helium tube. Measurement was based on absorption lines between 3889 A and Hγ, and reduction followed Mayall (**1**)

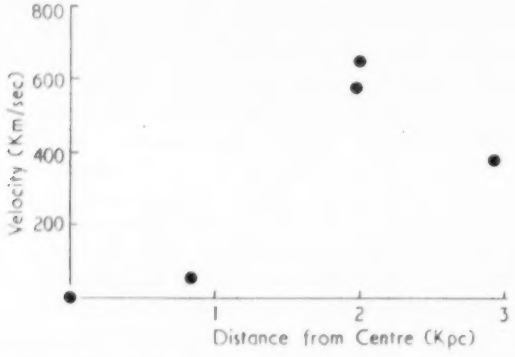
The results were as follows:

Region No.	Date (1956)	Exposure (hours)	Velocity (km/sec)
1	Sept. 5–6	6·4	− 68
2	Aug. 31–Sept. 1	7·3	+ 304
3	Aug. 30–31	7·5	− 115
7	Sept. 3–4	7·4	+ 354
8	Sept. 6–7	7·6	− 398

The velocities are reduced to the Sun. The regions are identified in Plate 10(*d*). Region W is that for which a velocity of −72 km/sec was derived at Mount Wilson (**1**). The region is described in that paper as "2′·7*n* of nucleus" which is not exactly correct for region W, but there seems to be no other feature which can be identified with the Mount Wilson region.

The spiral arm system of NGC 253 can readily be made out on the photographs. The nebula has been classified as Sc, but it seems possible, especially from a consideration of the plate taken in the red, that designation as a barred system might be more appropriate. It will be seen that the velocities given above imply that the spiral arms lag or trail in the rotation of the system.

Although the velocity measures given above can probably, at best, be trusted only to the nearest 50 km/sec, since the observations are near the limit of what is attainable with the present facilities and rather rudimentary guiding technique, there seems little room for doubt that the sense of rotation is now determined. Accepting the limitations of the results, it seems nevertheless worthwhile to push the discussion a little further, since the prospects of repetition or improvement of the observations are, at present, rather remote.

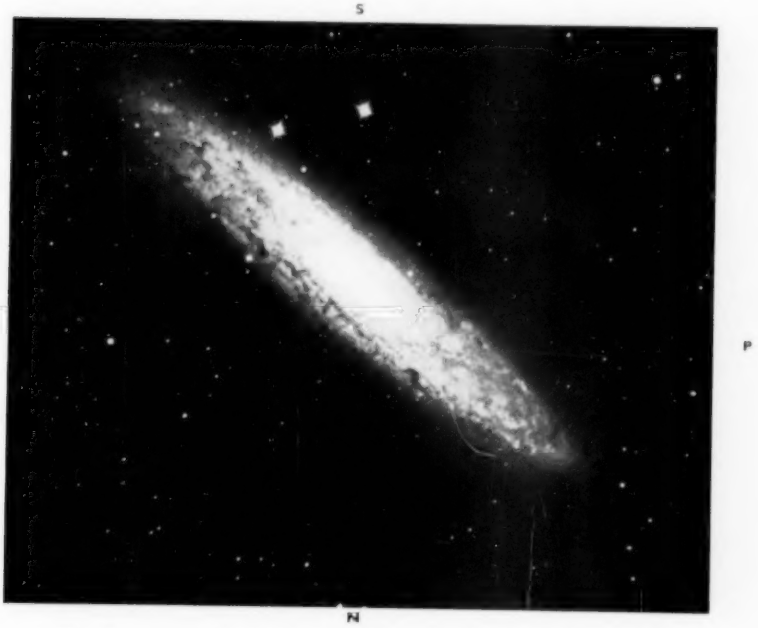


From the proportions of the system it may be inferred that the line of sight is inclined to the normal to the plane of the system at an angle given approximately as $\cos^{-1}(2/11)$. If we make the following assumptions—all doubtful—(i) that the observed velocities reflect only the local circular velocity and the bodily translation of the system, (ii) that the nuclear radial velocity is −68 km/sec, (iii) that all the measured features lie in the median plane of the system, and (iv) that the distance of the system is that given by Allen (**2**), viz. 1300 kpc, then we can derive points on the curve of circular velocity against distance. The geometry is obvious. The results are:

Region No.	Distance from centre (seconds)	(kpc)	Circular Vel. (km/sec)
1	0	0	0
2	325	2·05	598
3	134	0·84	54
7	331	2·09	667
8	470	2·96	375

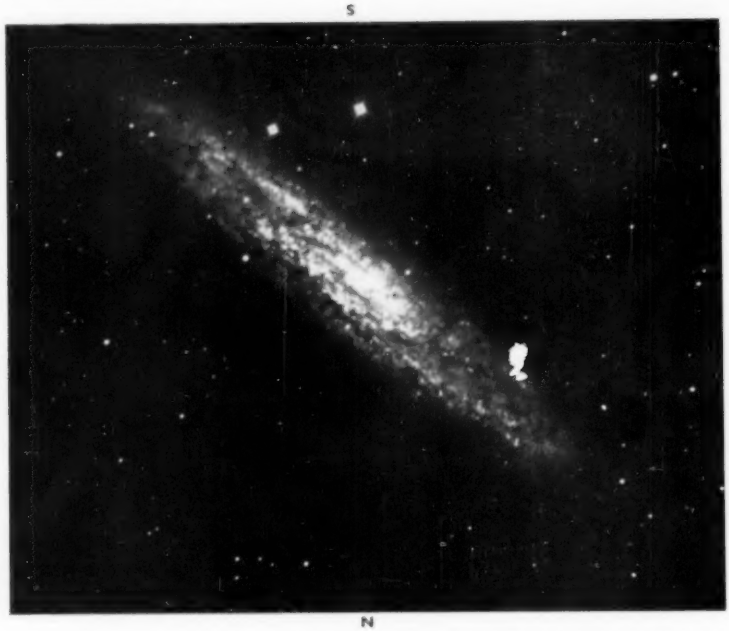
These results, given for what they are worth, are illustrated in the above figure.

(a) *Date:* 1953 *November* 9. *Emulsion:* 103a-E. *Exposure:* 150m. *Filter:* Ilford 205 (Narrow cut tricolour red, transmits beyond 6000 A and freely beyond 6400 A). *Scale:* 5″·7/mm.

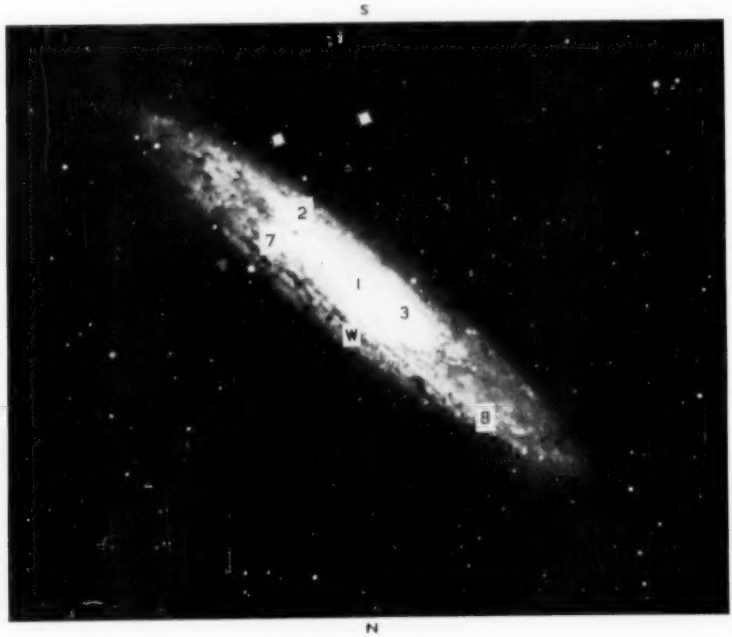


(b) *Date:* 1953 *November* 9-10. *Emulsion:* 103a-E. *Exposure:* 90m. *Filter:* Ilford 623 (Blue Green 4600 A–5450A). *Scale:* 5″·7/mm.

David S. Evans, The sense of rotation of NGC 253.



(c) *Date:* 1953 *November* 10. *Emulsion:* 103a-O. *Exposure:* 60m. *No Filter. Scale:* 5″·7/mm.



(d) *Identification of regions.*

I am deeply indebted to Dr H. C. Arp for his cooperation in guiding some of the spectra.

Royal Observatory,
Cape of Good Hope:
1956 *October* 10.

References

(1) M. L. Humason, N. U. Mayall and A. R. Sandage, *A. J.*, **61**, 97, 1956.
(2) C. W. Allen, " Astrophysical Quantities ", London, 1955.

VISUAL HORIZONS IN WORLD-MODELS

W. Rindler

(Received 1956 November 23)

Summary

This paper seeks to effect a unification and generalization of various particular results on visual horizons scattered in the literature. A horizon is here defined as a frontier between things observable and things unobservable. Two quite different types of horizon exist which are here termed event-horizons and particle-horizons. These are discussed in detail and illustrated by examples and diagrams. The examples include well-known model-universes which exhibit one or the other type of horizon, both types at once, or no horizon. Proper distance and cosmic time are adopted as the main variables, and the analysis is based on the Robertson–Walker form of the line element and therefore applies to all cosmological theories using a homogeneous and isotropic substratum.

---

1. *Introduction.*—There has recently occurred a renewed interest in the definition and properties of horizons in world-models, as witnessed by a prolonged correspondence in *The Observatory* (**1**) and subsequently in *Nature* (**2**). This discussion was chiefly devoted to the horizon in the de Sitter space-time associated with the model of the steady state theory, but the horizon concept has applications to many other models besides, and cosmologists since de Sitter and Eddington have concerned themselves with it. The above-mentioned correspondence drew attention to the lack of general agreement on even the definition of horizons. Different writers appear to use different horizon concepts, or different descriptions of the same concept, without always making the necessary distinctions. Furthermore, the meanings of many phrases used in discussions of horizons such as, for example, "all particles on one side of the horizon", "crossing the horizon with the speed of light", "in the observer's finite experience", etc. evidently depend critically on the definitions of time and distance whose diversity is notorious. A statement meaningful and valid on one interpretation can be meaningless or even false on another, unintended, interpretation. Consequently there can and did arise certain apparent "paradoxes". These facts seem to indicate the desirability for clear definitions to be laid down and accepted. That there is, moreover, room for a long overdue general study of horizons in cosmology was pointed out by H. Bondi and T. Gold (**3**) in the course of the *Observatory* correspondence.

In the present paper I seek to supply these needs. The main argument applies to all homogeneous and isotropic model-universes, being based on the Robertson–Walker form of the line element, but certain well-known particular models are singled out for illustrative purposes. As main variables I have chosen proper distance and cosmic time. In terms of these, horizons turn out to be loci of ordinary points, not singularities; in fact, it is now well-known that by means of these variables it is possible to represent the whole of space-time

regularly, in spite of what Eddington wrote in 1923*; the description of particles or events beyond the horizon presents no difficulties; and some apparent paradoxes disappear almost automatically. For the sake of clarity a small number of diagrams have also been designed. Their purpose is, however, purely schematic and no attempt was made to draw them to scale.

We shall define a horizon as *a frontier between things observable and things unobservable*. (The vague term *things* is here used deliberately.) There are then two quite different horizon concepts in cosmology which satisfy our definition and to which cosmologists have at various times devoted their attention. The first, which I shall call an *event-horizon*, is exemplified by the de Sitter model-universe. It may be defined as follows: *An event-horizon, for a given fundamental observer A, is a (hyper-) surface in space-time which divides all events into two non-empty classes: those that have been, are, or will be observable by A, and those that are forever outside A's possible powers of observation.* It was this horizon, with particular reference to the steady state theory, which formed the subject of the above-mentioned correspondence. Earlier references to it were made, among others, by Eddington (**5**) in connection with the de Sitter model (it was then called the "mass-horizon") and by E. A. Milne and G. J. Whitrow (**6**) in connection with Page's "uniformly accelerating" equivalence. A very interesting discussion of the event-horizon in the de Sitter model was recently given by Schrödinger (**7**) who greatly developed the geometric technique used by Eddington.

The other type of horizon, which I shall call a *particle-horizon*†, is exemplified by the Einstein–de Sitter model-universe. It may be defined as follows: *A particle-horizon, for any given fundamental observer A and cosmic instant t_0 is a surface in the instantaneous 3-space $t = t_0$, which divides all fundamental particles into two non-empty classes: those that have already been observable by A at time t_0 and those that have not.* W. H. McCrea (**8**) and G. C. McVittie (**9**) have previously discussed the particle-horizon from the point of view of its providing an empirical procedure for the determination of the rate of expansion of the universe. E. A. Milne (**10**) discussed it as an example of the absurdities inherent, in his opinion, in some of the cosmological models of General Relativity.

It must be pointed out that there are models, for example the well-known Lemaître model of General Relativity, which actually possess both types of horizon. Others, like Milne's "uniformly expanding" model possess neither type‡.

The general mechanism of horizons can be simply illustrated by means of the well-known "reduced" model which represents the universe as an expanding balloon. (This, of course, is a representation of a closed universe, but the argument for open universes is similar.) The fundamental particles can be represented by black dots distributed uniformly on the balloon. One particular dot may be marked specially so as to represent a given fundamental particle-observer *A*. Photons can be represented by red dots moving over the balloon along great

* "If de Sitter's form for an empty world is right it is impossible to find any coordinate system which represents the whole of real space-time regularly." (**4**). Our statement above includes de Sitter's form.

† It will be understood that whenever we speak of particles in this context we always mean fundamental particles, i.e. the representations of the nebulae in the world-model.

‡ What is commonly referred to as the horizon in that model is no true horizon in the sense of dividing things observable from things unobservable. The horizon of Milne's model is an accident of a particular choice of coordinates and can be "transformed away". See Section 10.

circle paths and always at constant speed relative to the material of the balloon. An event-horizon will exist for A, and similarly for all other fundamental observers, in models where the rate of expansion is and remains sufficiently great for some of the red dots moving on great circles towards A never to reach A. As Eddington put it, light is then like a runner on an expanding track, with the winning post receding faster than he can run. On the other hand, a particle-horizon will occur if, for example, the balloon is blown up from a volume approximating to a point at an initial rate exceeding the speed of the red dots, so that a finite time will elapse before any given one of them can reach A. *None* will reach A unless the rate of inflation decreases from its initial value. *Some* will never reach A if the rate of inflation, after first decreasing, increases again suitably. This is what happens in a model with both types of horizon.

A full discussion of these phenomena can, of course, not be given without a mathematical formulation, to which we now proceed.

2. *Mathematical preliminaries.*—As the mathematical basis of our discussion we take the Robertson–Walker line element,

$$ds^2 = c^2\,dt^2 - R^2(t)\left\{\frac{dr^2 + r^2(d\theta^2 + \sin^2\theta\,d\phi^2)}{(1 + kr^2/4)^2}\right\}. \tag{1}$$

In fact, the present paper may in a sense be regarded as a study of some of the properties of this metric. It has been shown to be applicable to all homogeneous and isotropic model-universes (**11**, **12**). Since practically all modern cosmological theories employ such model-universes our discussion will be of sufficient generality. The metric (1) has the following significance: (i) t is a cosmic time coordinate; (ii) θ, ϕ are the usual angular measurements made at the spatial origin $r = 0$ (which can be identified with any fundamental particle); (iii) the world-lines of the fundamental particles are the geodesics r, θ, ϕ = constant, whence r is a "co-moving" radial coordinate; (iv) light-tracks correspond to the null geodesics of the metric and, in particular, light-tracks through the spatial origin have the equations θ, ϕ = constant and

$$\frac{c\,dt}{R(t)} = \pm\frac{dr}{1 + kr^2/4}, \tag{2}$$

the positive sign evidently being required for light travelling away from the origin and the negative sign for light travelling towards it. Additional hypotheses are needed before a particular form can be assigned to the scale function $R(t)$ and a particular value (0, 1 or -1) to the curvature index k, and these are supplied by the various cosmological theories. In any theory which adopts the convention of the constancy of the local speed of light (and which consequently has interdependent time- and distance-scales) that speed is identified with the constant c in (1).

We shall employ proper distance and cosmic time as the main variables throughout this discussion. Due care must be exercised, especially in connection with the state of the universe shortly after the creation event in models where such an event occurs, not to identify these conventional variables too readily with "ordinary" distance and time.

If we introduce a function $\sigma(r)$, which may be regarded as an alternative "co-moving" radial coordinate, by the equation

$$\sigma(r) = \int_0^r \frac{dr}{1 + kr^2/4}, \tag{3}$$

then the proper distance, l, at time t_1 between the spatial origin and a fundamental particle with coordinate r_1 is given by $l = R(t_1)\sigma(r_1)$. (As is well-known, in theories with a conventionally constant speed of light this distance can be interpreted as the sum of the infinitesimal distance measurements made at the world-moment t_1 by a chain of intermediate fundamental observers situated along the space-geodesic joining the particle to the observer. In other theories l must be regarded as purely conventional.) The equation of motion of this fundamental particle is therefore

$$l = R(t)\sigma(r_1). \tag{4}$$

To obtain the equation of motion of a photon emitted at time t_1 at a fundamental particle with coordinate r_1 in the direction towards the origin, we integrate equation (2), choosing the negative sign, between the time and radial coordinate of emission and the current time and radial coordinate, t and r, and find

$$\sigma(r) = \sigma(r_1) - \int_{t_1}^{t} \frac{c\,dt}{R(t)},$$

which, on multiplication by $R(t)$, gives the required equation of motion,

$$l = R(t)\left\{\sigma(r_1) - \int_{t_1}^{t} \frac{c\,dt}{R(t)}\right\}. \tag{5}$$

In the sequel we shall want to assume that $\sigma(r)$ can take all positive values. Reference to (3) shows that this is evidently so when $k = 0$ or -1 (in the latter case the radial coordinate r is restricted by $r < 2$). When $k = 1$, $\sigma(r)$ as originally defined is restricted by $\sigma < \pi$ for all finite values of r, and thus there appears to exist a boundary to the fundamental particles at $l = \pi R(t)$. In fact this boundary is artificial and merely due to the particular definition of the coordinate r which in this case makes $r = \infty$ correspond to the antipole of the origin in the instantaneous spherical 3-spaces $t = \text{constant}$. There are, of course, particles beyond this pole on any line of vision. Moreover every particle occupies infinitely many positions on the line of vision, the latter being a closed curve of proper length $2\pi R(t)$. Now it is easily verified, and indeed it follows from the physical significance of proper distance, that σ is an additive coordinate in the sense that the σ of a particle C relative to the origin-particle A equals the σ of an intermediate particle B relative to A plus the σ of C relative to B. We therefore extend the definition of σ to particles beyond the pole by taking the pole as auxiliary origin, evaluating σ from there, and adding π. This process can be continued indefinitely. Analogy with an ordinary circle makes the idea quite clear: σ corresponds to the angular coordinate at the centre. Thus to each σ, however large, there corresponds a particle on the line of vision and each particle corresponds to infinitely many values of σ, all differing by multiples of 2π. A little care must therefore be exercised in the interpretation of horizons in these models, as will be seen in Section 4.

3. *The event-horizon.*—The necessary and sufficient condition for an event-horizon to exist in a given model is that the integral $\int^{\infty} \frac{dt}{R(t)}$ converge to a finite limit. For then, as reference to (5) shows, there exists at any given time t_0 and on any given line of vision a particle determined by

$$\sigma = \sigma_0 = \int_{t_0}^{\infty} \frac{c\,dt}{R(t)} \tag{6}$$

such that a photon emitted at that particle towards the origin at time t_0 reaches the origin at time $t = \infty$, i.e. in the infinite future. Photons emitted at the same time, t_0, from all farther particles ($\sigma > \sigma_0$) do not reach the origin at all (l never vanishes) whereas photons emitted at t_0 from all nearer particles ($\sigma < \sigma_0$) reach the origin at some finite time. We shall call the position of the critical particle at time t_0 the *horizon-point* at t_0 on that particular line of vision. If we multiply (6) by $R(t)$ and discard the particularizing suffix we obtain the equation of motion of the horizon-point.

$$l = R(t) \int_t^\infty \frac{c\,dt}{R(t)}. \tag{7}$$

Reference to (5) shows that this is actually the equation of motion of a photon, since (5) reduces to (7) if we put $\sigma(r_1) = \int_{t_1}^\infty \frac{c\,dt}{R(t)}$. Thus the horizon-point on any given line of vision may be identified with one particular photon travelling towards the origin. It follows that the event-horizon, which we now define as the aggregate of all horizon-points, is a closed light-front travelling towards the origin, and at proper distance from it given by (7).

It is perhaps useful to retain a dual picture of the event-horizon: (i) as a surface (geodesic sphere) in the observer's instantaneous 3-spaces $t = \text{constant}$, whose proper radius may or may not change with time, depending on the model, and (ii) as the corresponding hypersurface (a pseudo cylinder) in space-time. Where a distinction is necessary, the former might be called the space event-horizon, and the latter the space-time event-horizon.

Events occurring beyond this horizon are evidently for ever outside the possible powers of observation of the origin-observer A. On the other hand, *particles that have at some time been visible to A remain so for ever*. For, by equation (5), light emitted at time t_1 from a particle with coordinate r_1 reaches the origin at time t given by

$$\sigma(r_1) = \int_{t_1}^t \frac{c\,dt}{R(t)}.$$

If for some particular value $t = t_0$ this equation for t_1 furnishes a solution then it furnishes a solution for any $t > t_0$. Thus there is then a signal from the particle that arrives at the origin at any specified later time t. As $t \to \infty$, t_1 approaches a limiting value which evidently is the time at which the particle crosses A's horizon. Hence, although no particle can "pass out of view", its history as observed by A becomes more and more dilated, the event of its crossing the horizon being visible to A only in the infinite future, and all events at the particle after that event will never be visible to A. Again, *all fundamental particles other than A itself that are at some time within A's event-horizon, if such a horizon exists, must eventually pass beyond this horizon at the speed of light as measured locally*. These statements follow easily from equations (4) and (7), but they can also be seen in the following way. Consider any line of vision originating from a particle A and let B be another particle on this line, *within* A's horizon at some given time t. Then the proper distance of A's horizon-point P from A is the same as that of B's horizon-point Q from B, both distances being given by (7). Thus Q lies beyond P, whence P is a photon within B's horizon moving towards B, and will therefore reach B at a finite time. Moreover, like all other photons, it will reach B with the speed of light. This means that B will pass A's horizon-point, and pass it with the speed of light.

We may note that a contracting model-universe cannot have an event-horizon, since for the convergence of $\int^{\infty} \frac{dt}{R(t)}$ it is necessary that $R(t)$ increase to infinity. Models oscillating between singular states must logically be treated as existing only for a finite time (the time of one oscillation) and so the concept of an event-horizon would appear to be irrelevant to them.

4. *Examples of models possessing an event-horizon.*—As an important first example we consider the de Sitter model-universe. In this model $R(t)=e^{t/T}$, where T is the reciprocal of Hubble's coefficient and has the dimensions of time, and $k=0$. The model evidently satisfies the condition given at the beginning of Section 3 for the existence of an event-horizon. From (4) the equations of motion of fundamental particles in the de Sitter model are

$$l=\sigma(r_1)e^{t/T}, \tag{8}$$

which correspond to similar exponential curves on an l–t diagram (see Fig. 1) referring to a single semi-infinite line of vision. All these curves have the t-axis as asymptote. From (5) we find that the equations of motion of photons moving towards the origin are given by

$$l=cT+\gamma e^{t/T}, \quad \gamma=\sigma(r_1)-cTe^{-t_0/T}. \tag{9}$$

These equations also correspond to similar exponential curves, but they have the line $l=cT$ for asymptote. They lie above or below this asymptote according as γ is positive or negative. (There is no contradication in the fact that a photon travels towards A, yet its proper distance from A increases: cf. Eddington's runner mentioned in the Introduction.) It is evident that the line $l=cT$ represents

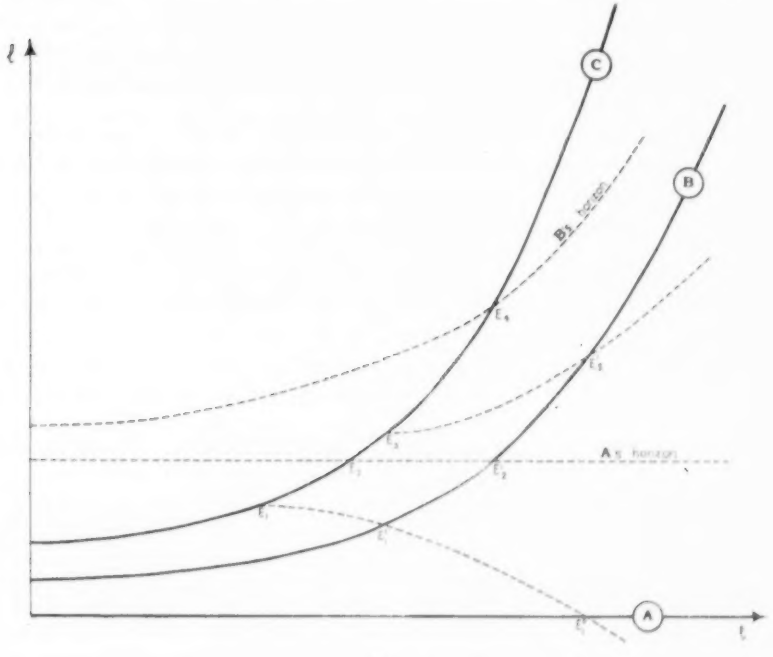


FIG. 1.

the event-horizon, and this naturally agrees with (7). In Fig. 1 only three particle paths have been drawn (full lines)—those of the origin-particle A, and of two other particles, B and C. The stippled curves represent light-paths corresponding to photons travelling towards the origin and two of these, in particular, are the paths of the horizon-points of A and B. (It can easily be verified, and indeed it is obvious from symmetry considerations, that light-paths directed *away* from the origin-particle correspond to similar exponential curves with the line $l=-cT$ for asymptote. However, none of these has been marked in the diagram.) E_1, E_2, E_3, E_4 are events, in that order, at C. E_1 is observed at B at E_1' and at A at E_1''. E_2 is the event of C's crossing A's horizon, and this is observed at B at E_2', but the corresponding event E_2'' at A occurs in the infinite future. E_3, which occurs beyond A's horizon but within B's, will be observed at B at E_3' but never at A. Events after E_4 (the crossing of B's horizon by C) are unobservable also at B.

The same diagram can be used to illustrate the significance of a horizon in models with positive space-curvature ($k=1$) which present some complications as was pointed out at the end of Section 2. Let us consider a model which has $R(t)=e^{t/T}$ and $k=1$. The above calculations are equally valid for this model, only the interpretation differs somewhat. We shall now suppose that A, B and C are representations of the same particle (at $\sigma=0$, 2π, 4π respectively). Thus, for example, the segment E_1, E_1', E_1'' represents a light-track going twice round the universe. If E_1' and E_3 are the same event, then so are E_1'' and E_3'. E_2' and E_4 are evidently identical. All events occurring before t_2' (the time of E_2') are effectively observable at A, that is, at A or B or C etc. At time t_2' each particle crosses its own horizon for the last time. Before t_2' the expansion of the model is such that each photon is able to make at least one complete circuit of the universe, but for photons emitted after t_2' no more complete circuits are possible. The domain of unobservable events for A is thus not given by $l>cT$ alone but by $l>cT$ and $t>t_2'$.*

As a second example we consider the class of models characterised by $R(t)=at^n$, where a is a constant. This class contains several well-known models. For example, the Einstein–de Sitter model with $n=2/3$, Dirac's model with $n=1/3$, Milne's "uniformly expanding" model with $n=1$ and Page's "uniformly accelerating" model with $n=2$ (**13**). Evidently the necessary and sufficient condition for this type of model to have an event-horizon is $n>1$. Consequently the Einstein–de Sitter model, Dirac's model and Milne's model do not have an event-horizon. From (7) the equation of the horizon, when it exists, is

$$l=\frac{ct}{n-1}, \tag{10}$$

which shows that the horizon expands uniformly in all cases. For illustration we may choose the typical case of Page's model which has $R(t)=at^2$. From (4) the particle paths are $l=a\sigma(r_1)t^2$ and from (5) the light-paths are $l=\gamma t^2+ct$, where $\gamma=a(\sigma(r_1)-c/t_1)$. On an l–t diagram the former are parabolas touching the t-axis at the origin, whereas the latter are parabolas passing through the origin with gradient c and curving away from or towards the t-axis according as γ is

* Events unobservable from this one direction may, however, be observable from the opposite direction. Only after A's antipode $\bar{A}(\sigma=\pi)$ crosses A's horizon for the last time, does there develop around $\bar{A}$ an ever-growing region whose events are truly unobservable by A.

positive or negative. The line $l = ct$ corresponding to $\gamma = 0$ divides these two classes and thus evidently represents the event-horizon of the origin-observer, which is in agreement with (10).

5. *The particle-horizon.*—The necessary and sufficient condition for a particle-horizon to exist in a given model is the convergence of the integral $\int_0 \frac{dt}{R(t)}$, or of $\int_{-\infty} \frac{dt}{R(t)}$ in cases where the definition of $R(t)$ extends to negatively unbounded values of t. The analysis is quite similar in the two cases and we therefore assume the former condition to obtain, since the latter implies the less probable situation of a universe contracting from a state of infinite expansion. It then follows from (5) that at any given time t_0 all those particles for which $\sigma > \int_0^{t_0} \frac{c\,dt}{R(t)}$ have not yet been observed by the origin-observer A, whereas all others have. Hence the surface

$$\sigma = \int_0^{t_0} \frac{c\,dt}{R(t)} = \phi(t_0), \tag{11}$$

where ϕ is defined by this equation, divides all particles into two non-empty classes: those that have been observed by A at or before t_0, and those that have not. This surface we call A's particle-horizon at time t_0. It is evidently a geodesic sphere in the observer's instantaneous 3-space $t = t_0$ and its radius may or may not change with time, depending on the model. Occasionally we shall wish to regard the particle-horizon as the section $t = t_0$ of the hypersurface

$$\sigma = \phi(t) \tag{12}$$

in space-time, which might be called A's space-time particle-horizon.

As $R(t)$ is by its nature positive and finite, $\phi(t)$ is an increasing function of t when it exists, and thus, by (11), *more and more particles become visible to A as time goes on*, since σ is a "co-moving" coordinate. Whether *every* particle eventually becomes visible to A or not depends on whether $\phi(t)$ tends to infinity with t or not. If $\phi(t)$ approaches a finite limit then all those particles for which $\sigma > \int_0^{\infty} \frac{c\,dt}{R(t)}$ are entirely outside A's possible powers of observation and the model evidently possesses an event-horizon in addition to its particle-horizon. In any case, as before, a particle once seen remains for ever visible. Each particle is first seen at its "birth" or, in other words, the first signal received from the particle in question was emitted there at $t = 0$ (or $t = -\infty$, as the case may be). However, in models which postulate a unique creation event in the finite past ($R(0) = 0$) each particle comes into view with an infinite Doppler ratio. For the latter is given by the well-known formula (**14**)

$$\frac{\lambda_0}{\lambda_1} = \frac{R(t_0)}{R(t_1)},$$

where λ_1, t_1 and λ_0, t_0 are the wave-lengths and times of emission and reception, respectively, of the light-signal whereby the particle is observed, and in our case $R(t_1) = R(0) = 0$. This would, in practice, prevent the observation of the actual creation process, which was one of the objections Milne (*loc. cit.*) made to models possessing a particle-horizon. On the other hand, in models which contract from infinite extension in the infinite past and possess a particle-horizon (e.g. a model with $R(t) = e^{-t/T}$), particles first come into view with zero wave-length!

The unique creation event (point-creation) is a singularity in all models in which it occurs, in the following sense: there is one particle world-line, and only one, through each event except the creation event, through which there passes an infinity of such world-lines. But in models with point-creation and a particle-horizon, the creation event is a singularity in yet another sense: each event determines a unique light-cone with vertex at that event except the creation event at which there originates an infinity of different light-cones, one corresponding to each particle. For consider the equation

$$l = R(t)\left\{\sigma(r_1) \pm \int_{t_1}^{t} \frac{c\,dt}{R(t)}\right\}, \tag{13}$$

which is obtained from (2) in the same way as (5) but with the retention of both signs. This is an equation of the light-cone whose vertex is at a particle characterized by r_1 at time t_1. Now if $\int_0 \frac{dt}{R(t)}$ is divergent, the light-cone of the creation event ($t_1 = 0$) has equation $t = 0$, the same for all particles. But if this integral is convergent the light-cone of the creation event is evidently not unique but depends on which particle emits the light.

Let us consider now the $l-t$ equation of the space-time particle-horizon. We obtain this by multiplying (12) by $R(t)$ and using (11),

$$l = R(t)\int_0^t \frac{c\,dt}{R(t)}. \tag{14}$$

Comparison of (14) with (13) shows that *the space-time particle-horizon of any observer is the boundary of his creation-light-cone*. Thus a particle B becomes visible to the particle-observer A the moment it enters A's creation-light-cone. This is not surprising, for, by the symmetry that obtains between the two particles, A at that same instant enters B's creation-light-cone. Fig. 2 illustrates this situation.

In the same way that we identified the event-horizon with a geodeso-spherical light-front converging on the observer and reaching him at time $t = \infty$, so from our present discussion of the particle-horizon we see that this latter can be identified with a geodeso-spherical light-front diverging from and emitted by the observer at time $t = 0$ (or at $t = -\infty$, as the case may be), being the section $t =$ constant of the space-time particle-horizon. It is therefore evident that fundamental particles entering this horizon do so at the speed of light, as measured locally. This is why the particle-horizon has sometimes been described, perhaps misleadingly, as expanding with the speed of light. Its proper distance from the observer, given by (14) or the corresponding formula with an integral from $-\infty$ to t, may remain constant or even decrease.

6. *Examples of models possessing a particle-horizon.*—Of the class of models characterized by $R(t) = at^n$, where a is a constant, evidently only those models which have $n < 1$ possess a particle-horizon which, by (14), has equation

$$l = \frac{ct}{1-n}. \tag{15}$$

Thus, in particular, Milne's uniformly expanding model ($n = 1$) and Page's uniformly accelerated model ($n = 2$) do not possess a particle-horizon, whereas Dirac's model ($n = 1/3$) and the Einstein–de Sitter model ($n = 2/3$) do possess one. I shall choose this latter model for illustration since it is typical. From (4)

we obtain the equations of the particle paths, $l = a\sigma(r_1)t^{2/3}$, and from (13) we find the equations of the paths of photons emitted at an arbitrary particle with coordinate r_1 at the creation event ($t_1 = 0$) and directed towards and away from the origin,

$$l = t^{2/3}\{a\sigma(r_1) \pm 3ct^{1/3}\}. \tag{16}$$

In Fig. 2 the world-lines of only two particles (one of which is taken to be the origin-particle) are shown, together with their corresponding creation-light-cones. We note that these cones have no overlap near their common vertex, and this is typical for models with a particle-horizon. At time t_0 each of the two particles enters the creation-light-cone of the other, and thus they become visible to each other for the first time.

It may be of interest to quote Milne's description (*loc. cit.*) of the horizon in the Einstein–de Sitter model: "The frontier... of observability moves onward with the speed of light; it contains always the particles just being created and it leaves in its wake a spray of decelerating newly created particles." Although Milne's deliberately absurd deductions can to some extent be blamed on the subjective coordinates he used, certain physical difficulties seem to be inherent in models possessing a particle-horizon: if the model postulates point-creation we have material particles initially separating at speeds exceeding those of photons; if the model postulates collapse from infinite rarification there is the extraordinary phenomenon, noted above, of particles coming into view with zero wave-length; lastly, models postulating a beginning at a finite epoch with finite separation between particles, whilst possessing neither of these crass difficulties, seem hardly less artificial.

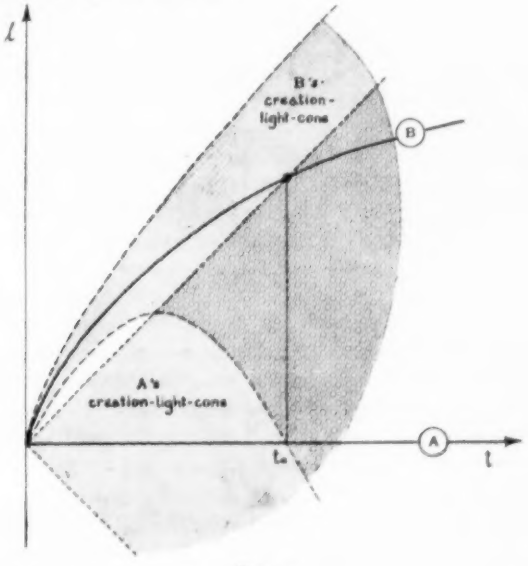


FIG. 2.

As an example of a model-universe possessing both a particle-horizon *and* an event-horizon I shall choose that General Relativity model which is characterized by $k = 0$ and a positive cosmological constant, and has

$$R(t) = a(\cosh bt - 1)^{1/3},$$

where a and b are constants (**15**). This model is somewhat similar in general behaviour to the better known Lemaître model, but has the advantage, for illustrative purposes, of possessing a simple functional form of $R(t)$. We need not here enter into the exact calculations for the light-paths, but we note that $R \sim t^{2/3}$ for small values of t, and $R \sim \exp(bt/3)$ for large values of t, whence the model is seen to approximate asymptotically to the Einstein–de Sitter model for small values of t and to the de Sitter model for large values of t. The properties of this model are illustrated schematically in Fig. 3. The origin-observer is denoted by A. B is an observer on a typical particle which becomes visible to A at creation-time t_1 (when A and B enter each other's creation-light-cones) and which passes beyond A's event-horizon at time t_2, so that events at B after t_2 are outside A's possible powers of observation. C is the critical particle which becomes visible to A only at $t = \infty$. C's creation-light-track towards A is that of the unique photon which reaches A at $t = \infty$, and which we have already identified with A's event-horizon. And in the same way that A approaches asymptotically the boundary of C's creation-light-cone, so C approaches that of A's creation-light-cone. Evidently all particles beyond C are entirely outside A's cognizance. In the diagram only the portions near the vertices of the creation-light-cones have been shaded in.

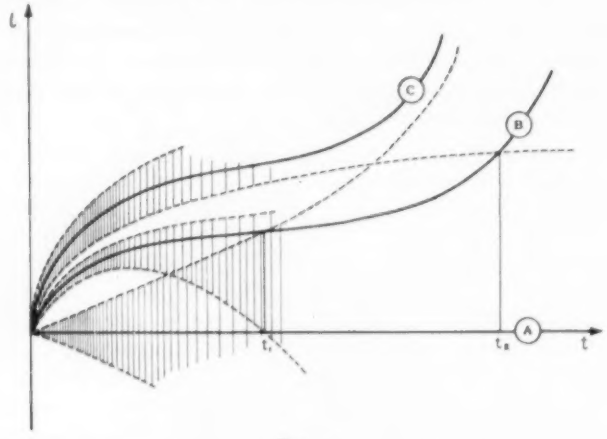


FIG. 3.

We may note that the existence of a critical particle with properties analogous to those of C above, one on each line of vision of each fundamental observer, is a general feature of all models possessing both types of horizon. From equations (4) and (5) we easily find that the σ-coordinate of this particle relative to the observer is given by

$$\sigma_c = \int_0^{\infty} \frac{c\,dt}{R(t)} \quad \text{(or} \quad \sigma_c = \int_{-\infty}^{\infty} \frac{c\,dt}{R(t)}\text{, as the case may be).}$$

7. *Horizons for non-fundamental observers.*—In connection with the de Sitter model Eddington (**16**) made the assertion that "events before $t = -\infty$ (i.e. events outside the event-horizon) may produce consequences in the neighbourhood of the observer and he might even see them happening..." Schrödinger (**17**) in his recent book concurred with this assertion and amplified it to some

extent. Evidently the observer contemplated by these authors cannot be a fundamental observer. In defining the event-horizon of a particular fundamental observer in an arbitrary model we assumed that the observer in question remained attached to his original fundamental particle. If we discard this restriction the class of events observable by him is naturally increased, and we shall now briefly examine this new situation.

By a generally accepted relativistic principle an observer can theoretically reach and be present at any preassigned event within his forecone. This is merely equivalent to saying that matter can attain any velocity short of the velocity of light, both velocities being measured locally and with respect to the substratum. Thus if at some instant t_0 an observer can detach himself from his fundamental particle and journey with sufficient energy into space, thereby ceasing to be a fundamental observer at least for a certain period, all events within his forecone at t_0 become attainable to him and all events whose forecones intersect his own become observable by him. We shall now prove the following statements, in so far as they have not already been proved:

(i) If a model possesses no event-horizon then all events are observable sooner or later by any fundamental observer.

(ii) If a model possesses an event-horizon but no particle-horizon then any observer can be present at any one preassigned event provided he is willing if necessary to detach himself from his original fundamental particle and provided also he does so soon enough. Nevertheless, if two events are preassigned it will, in general, be impossible for any one observer either to be present at, or ever to observe, both events.

(iii) If a model possesses an event-horizon *and* a particle-horizon then for any observer originally attached to a given fundamental particle there exists a class of events absolutely beyond his cognizance, no matter how he journeys through space. We may here speak of an ***absolute horizon.***

Statement (i) has been justified in Section 3. To prove the first part of (ii) we must show that any event E lies in at least one forecone of any given fundamental observer A. Let E occur at a fundamental particle B at cosmic time t_0. Then if E lies in none of A's forecones, A has not been observable at B at or before time t_0, and this contradicts our assumption that there is no particle-horizon. Thus E lies in one of A's forecones and the statement is proved. In support of the second part of (ii) we shall show that corresponding to any event E there exists a class S of events such that no event of S can be observed by any observer who ever observes E. Consider the closed light-front π converging on the fundamental particle B associated with E and at the time t_0 of E at proper distance $2\rho_0$ from B, where ρ_0 is the radius of the event-horizon at t_0. Any two photons emitted towards each other at t_0 a proper distance $2\rho_0$ apart will meet at $t=\infty$. Hence photons emitted at E meet π at $t=\infty$. This proves that the forecones of events occurring beyond π will not intersect the forecone of E. These events evidently constitute the class S. For an observer seeing an event must be situated in the forecone of that event and cannot ever get out of it again. We may note that, conversely, the forecones of all events occurring within π do intersect the forecone of E and hence any one of these events can be observed jointly with E.

Lastly we prove (iii). We remember that in any model possessing both types of horizon there exists on any line of vision of any given fundamental observer A

a critical particle C with coordinate σ_e which enters A's creation-light-cone at $t=\infty$ (cf. Section 6). Consider now the particle P with coordinate $2\sigma_e$ on the same line of vision. Clearly C enters P's creation-light-cone also at $t=\infty$ and therefore the creation-light-cones of A and P intersect at $t=\infty$. Thus it is impossible for A ever to observe the particle P at a finite time, even if he detaches himself from his original fundamental particle. He can therefore have absolutely no cognizance of events occurring at P and, *a fortiori*, of events occurring beyond P. All events occurring nearer than P can, on the other hand, be observed by A provided he can if necessary detach himself from his fundamental particle and journey within his creation-light-cone. Thus the path of P, which has equation

$$l=2\sigma_e R(t), \tag{17}$$

constitutes an absolute horizon.

8. *Effects of reversing the direction of time.*—We have noted the phenomenon that in certain models (those possessing a particle-horizon) fundamental particles suddenly come into the view of fundamental observers at a finite time. We have also noted that the reverse phenomenon, viz. particles disappearing from view at a finite time, cannot occur under any circumstances. A very loose argument might seem to suggest that if we reverse the direction of time in any model with a particle-horizon we ought to obtain a model in which particles disappear from view. Although this is by no means the case it is of some interest briefly to consider such time reversals, especially in view of the fact that in some models (e.g. all the cosmological models of General Relativity) the direction of time can be reversed without violating the hypotheses on which the model is constructed. In any case there is nothing to prevent us from contemplating the dual of any given model formed in this way. The one result that is of interest in this connection is that an event-horizon transforms into a particle-horizon and vice versa.

Let us first consider a model with a scale-function $R(t)$ defined over the whole range $-\infty<t<\infty$ and possessing an event-horizon whose equation will be, as we saw in (7),

$$l=R(t)\int_t^\infty \frac{c\,dt}{R(t)}. \tag{18}$$

The dual of this model has the scale-function $R(-t)$ and thus possesses a particle-horizon whose equation, by analogy with (14), is given by

$$l=R(-t)\int_{-\infty}^t \frac{c\,dt}{R(-t)}. \tag{19}$$

This equation also results if we change the free variable t in (18) into $-t$ and thus our assertion is proved. Let us also consider a model with point-creation in the finite past and a particle-horizon. On time reversal the point-creation event transforms into a point-annihilation event in the finite future. The particle-horizon transforms into an event-horizon in the sense that events occurring beyond it will not be observed in the finite stretch of time left to the observer before annihilation. With similar modifications our result applies in all cases.

We may also note that in the dual model all Doppler ratios are inverted. If a light-signal from an event E_1 to an event E_2 in the original model exhibits a Doppler ratio $D=\dfrac{\lambda_2}{\lambda_1}$, where λ_1 and λ_2 are the wave-lengths as measured at E_1

and E_2 respectively, then the corresponding signal from E_2 to E_1 in the dual model exhibits a Doppler ratio $1/D$. This is easily seen by considering two successive light-signals whose time difference is dt_1 at E_1 and dt_2 at E_2. Then the Doppler ratios in the original model and the dual are $\frac{dt_2}{dt_1}$ and $\frac{dt_1}{dt_2}$ respectively.

9. *General time transformations.*—It is one of the achievements of the theory of Kinematic Relativity that it brought into prominence the relativity of clock rates. No one clock rate is any longer regarded as absolute. Cosmological theories other than steady state theories must recognize the possibility that even different physical clocks, e.g. "atomic" clocks and "dynamic" clocks, though synchronous momentarily will not necessarily go on keeping the same time indefinitely. Again, Kinematic Relativity showed how it was possible to adopt a conventionally constant speed of light by linking the unit of distance to the unit of time. Thus the only standard unit required is the unit of time. If we change from one standard clock to another whose rate, let us assume, accelerates relative to the first, then the corresponding new unit of length decreases relative to the old one. Distances which in our first reckoning remained constant will now be regarded as increasing with time. Thus by means of a suitable time transformation it is possible to transform a static model-universe into an expanding one, or indeed to transform a model with any given rate of expansion into one with any other given rate of expansion. The detailed formulation and proof of this result for models described as "kinematic equivalences" is sometimes referred to as the "main theorem on equivalences" (**18**). I have shown elsewhere (**13**) that this theorem applies equally to models described in terms of the metric (1). The question arises whether, if no one clock rate can be regarded as absolute, all models are indeed equivalent to each other. This is an impression one sometimes gains from Milne's writings.

There are physical objections to this view: the clocks implicit in our empirical work on the expansion of the universe are atomic clocks and not arbitrary clocks (**13**) and the model we seek is unique. However, even theoretically there is a limitation to the formal equivalence of different models, and horizons provide the clue to this limitation. Eddington (**19**) considered an essentially similar problem and may even have anticipated Milne in his interrelation of the units of time and distance. Eddington concluded that a reduction of the de Sitter model to rest by means of a time transformation leads to an absurdity, namely the sudden arrest of all proper motions at a finite future time. We can now say that this result is characteristic of all models possessing an event-horizon. For the clocks that give rise to the static description of the model can be identified with rays of light oscillating between neighbouring fundamental particles, successive reflections defining equal units of time. We know that if the model possesses an event-horizon there will come a time when the signal can no longer pass from one of these particles to the other. Evidently each photon in the static description of the model moves with uniform speed towards a limiting particle and then stops dead. All other proper motions must cease at that instant also. Hence models with event-horizon when transformed to the static form exhibit properties startlingly different from those of genuinely static models which exist indefinitely. Evidently no model with an event-horizon can be equivalent to one without, and a similar remark holds for models with particle-horizons. Transformations of one of these types

of model into another are possible only partially, a fact already noted by Milne and Whitrow (**6**) in connection with Page's and Milne's models.*

10. *The boundary in Milne's model.*—Milne's model is often described in coordinates different from the cosmic time and proper distance employed here, namely as an expanding spherical aggregate of particles satisfying the postulates of homogeneity and isotropy and embedded in the space-time of the special theory of relativity. Each fundamental particle-observer may consider himself to be at the centre of the sphere, whose boundary expands with the speed of light and which therefore has radius ct at time t, in the coordinates of the special theory of relativity. It is well known that this description of the model is equivalent to its description by means of the metric (1) with $R(t)=ct$ and $k=-1$, and we saw that that model possesses neither an event-horizon nor a particle-horizon. Nevertheless, using the first description, it has seemed quite natural to Milne to call the boundary of the sphere a "horizon". What is the intrinsic significance of this "horizon" when we leave out of account the particular coordinates employed? It is simply that each observer considers the distances of fundamental particles from himself to possess an unattained upper bound. It appears to him at time t that this bound is at a distance $\frac{1}{2}ct$ from him (world-picture) but taking into account the finite velocity of light he calculates that the bound is actually at distance ct (world-map). In order to associate an intrinsic significance with the distance concept we note that in the special theory of relativity distance from parallax is identical with coordinate distance. Thus we can finally characterize the situation in Milne's model as follows: All fundamental particles are visible at all times and there exists at any time t a finite upper bound $\bar{P}(t)$ to the apparent distances by parallax of these particles. When we inquire whether this state of affairs is in any way peculiar to Milne's model we find that this is by no means the case.

The distance by parallax, P, is defined by the equation $P\,d\theta=dl$, where $d\theta$ and dl are measured at the observer and represent angular difference and distance respectively between two neighbouring rays diverging from the event in question. It can be shown (**20**) that in a model referred to the metric (1) the distance by parallax of an event occurring at coordinate r_1 and observed at time t_0 is given by

$$P=\frac{R(t_0)r_1}{1-\frac{1}{4}kr_1^2+R'(t_0)r_1/c}. \tag{20}$$

* The mathematical formulation is straightforward: to transform a model with metric (1) into another with the same curvature index k but a different scale-function, say $G(t)$ instead of $R(t)$, let us apply the clock transformation $t=g(T)$. The metric becomes

$$ds^2=(g'(T))^2\left\{c^2\,dT^2-\left(\frac{R(g(T))}{g'(T)}\right)^2 d\rho^2\right\},$$

where $d\rho^2$ is the $\{\ \}$ of (1). The conformality factor is discarded and the Robertson-Walker metric of the new model is $ds^2=c^2\,dT^2-G^2(T)\,d\rho^2$, provided $\frac{R(g(T))}{g'(T)}=G(T)$, which gives the required $g(T)$. This equation also yields the relation

$$\int_{t_0}^{t_1}\frac{dt}{R(t)}=\int_{T_0}^{T_1}\frac{dT}{G(T)}$$

for corresponding time intervals t_1-t_0, T_1-T_0. Suppose the R model possesses an event-horizon so that the left integral converges for $t_1\to\infty$. If the G model does not possess such a horizon evidently $t_1=\infty$ transforms into a *finite* time T_1 and greater values of T do not correspond to any values of t. The situation is similar when only one of the models possesses a particle-horizon.

In Milne's model we therefore have

$$P = \frac{ct_0 r_1}{(1+\frac{1}{2}r_1)^2}, \tag{21}$$

whence P approaches the value $\bar{P}(t_0) = \frac{1}{2}ct_0$ as r_1 approaches its limiting value of 2 (for the most distant particles), which agrees with our previous assertion.

Let us next consider the general class of models characterized by

$$R(t) = ct^n \quad (n>1), \quad k=-1.$$

From (20) we find that, for this class,

$$\bar{P}(t_0) = \frac{ct_0^n}{1+nt_0^{n-1}}, \tag{22}$$

and thus all these models possess a Milne type of boundary which, moreover, to a first approximation expands uniformly. When we consider a model of the above general class but with $n<1$ we find the situation somewhat changed: although there still exists a finite upper bound to the distances by parallax of all visible particles, not all particles are visible at any given time, as we saw in Section 6.

We may note, finally, that the de Sitter model ($R(t) = e^{t/T}$, $k=0$) also possesses a Milne type of boundary with constant radius $\bar{P} = cT$, as is easily verified.

11. *Acknowledgments.*—Certain portions of this work were first presented as a chapter of a Ph.D. thesis submitted by me to the University of London in April 1956. In this connection I wish to thank my supervisor, Dr G. J. Whitrow, and my external examiner, Professor H. Bondi, for their interest and discussion. I also wish to express my thanks to Mrs Margaret Henderson for the patience and skill with which she executed the diagrams.

Department of Mathematics,
Cornell University, Ithaca, New York, U.S.A.:
1956 *November.*

References

(1) *The Observatory*, **73**, 205, 1953; **74**, 36, 37, 172, 173, 1954.
(2) *Nature*, **175**, 68, 382, 808, 1955.
(3) *The Observatory*, **74**, 36, 1954.
(4) A. S. Eddington, *The Mathematical Theory of Relativity* (2nd edition), p. 166, Cambridge, 1924.
(5) A. S. Eddington, *loc. cit.*, pp. 157, 165.
(6) E. A. Milne and G. J. Whitrow, *Z. Astrophys*, **15**, 345, 1938.
(7) E. Schrödinger, *Expanding Universes*, sections 4, 5, 6, 8, Cambridge, 1956.
(8) W. H. McCrea, *Z. Astrophys.*, **9**, 290, 1935.
(9) G. C. McVittie, *Cosmological Theory*, p. 54, London, 1937.
(10) E. A. Milne, *Relativity, Gravitation and World Structure*, p. 327, Oxford, 1935.
(11) H. P. Robertson, *Astrophys. J.*, **82**, 284, 1935.
(12) A. G. Walker, *Proc. Lond. Math. Soc.* (2) **42**, 90, 1937.
(13) W. Rindler, *M.N.*, **116**, 335, 1956.
(14) H. Bondi, *Cosmology*, formula (10·18), Cambridge, 1952.
(15) H. Bondi, *loc. cit.*, pp. 82 and 104.
(16) A. S. Eddington, *loc. cit.*, p. 166.
(17) E. Schrödinger, *loc. cit.*, p. 21 *et seq.*
(18) E. A. Milne, *Kinematic Relativity*, section 16, Oxford, 1948.
(19) A. S. Eddington, *The Expanding Universe*, chapter III, section VI, Cambridge, 1932.
(20) W. H. McCrea, *loc. cit.*, formula (15).

THE GRAVITATIONAL FIELD OF A STAR IN THE EXPANDING UNIVERSE

C. Gilbert

(Received 1956 October 10)

Summary

A problem previously considered by Einstein and Straus is solved by the use of new boundary conditions proposed by O'Brien and Synge. The solution joins the Schwarzschild exterior solution of Einstein's field equations to the field generated by a uniform distribution of matter in the expanding universe. It is discussed in relation to previous work on the problem.

---

1. *Introduction.*—The influence of the expansion of space on the gravitational field in the neighbourhood of a star, was investigated by Einstein and Straus in a well-known work (**1**). In their approach to the problem they assumed that the star was situated *in vacuo* at the centre of a spherical hole in the otherwise uniform distribution of matter of an expanding universe. The field in the hole was assumed to be centrally symmetric and was joined to the field generated by the exterior matter, by assuming that the field tensors and their first derivatives with respect to the spatial coordinates, were continuous, at all times, at the boundary of the hole.

Einstein and Straus proved that a unique solution of the problem of determining the field in the neighbourhood of the star existed, without however solving their equations, and they also showed that the interior field could be transformed into the static Schwarzschild solution. They later pointed out (**2**) that the latter result was a consequence of a general theorem proved by Birkhoff (**3**), that every non-static centrally symmetric field can be transformed into the Schwarzschild solution. It should also be noticed that it follows from Birkhoff's theorem, that any solution of the Einstein–Straus problem which may be based on different boundary conditions from those stated above, can be transformed into the Schwarzschild solution.

It is of interest to solve the problem completely in order to determine the transformation from the "local" coordinates of the Schwarzschild solution to the "cosmic" coordinates of the line-element for the expanding universe. The solution also yields an interesting formula for the gravitational mass of the star, and is also of interest in its applications to the field of a globular cluster of stars, or of a spherical galaxy, in an expanding universe. The work of Qvist (**4**), Kustaanheimo (**5**) and Schücking (**6**) is concerned with completing the solution. Schücking obtains a formula in agreement with the results of this paper, for the case when the cosmological constant is zero. He considers only the continuity of the field tensors, and it is not therefore clear that his solution satisfies the boundary conditions used by Einstein and Straus.

In this paper a solution of the problem is obtained by using the boundary conditions recently proposed by O'Brien and Synge (**7**). Bonnor (**8, 9**) has discussed and used these boundary conditions in recent work, and considers them to be the best available at present.

2. *The boundary conditions for the Schwarzschild field.*—The field of the expanding universe may be defined by the metric

$$ds^2 = -R^2(f^{-2}dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\phi^2) + c^2dt^2, \tag{1}$$

where R is a function of t only, $f^2 = 1 - r^2/R_0^2$ and c is the velocity of light. R_0 is a constant which may be real, imaginary or infinite. We join this field at the 3-space $r = a$, subsequently referred to as Σ, on to a centrally symmetric field defined by the metric

$$ds^2 = -e^{\lambda}dr^2 - e^{\omega}(d\theta^2 + \sin^2\theta d\phi^2) + e^{\nu}dt^2, \tag{2}$$

where λ, ω, ν are functions of r, t.

The boundary conditions at Σ are assumed to be those given by O'Brien and Synge (**7**). If we denote by g_{mn} and T_{mn} the fundamental tensor and stress-energy tensor for (1) or (2), and write $x^m (m = 1, 2, 3, 4)$ for the coordinates r, θ, ϕ, t, these conditions are that

$$g_{mn},\ \partial g_{\mu\nu}/\partial x^1,\ T^1_n,\ g_{m\rho}T^{\rho}_n - g_{n\rho}T^{\rho}_m,$$

shall be continuous at Σ, Latin indices taking the values 1, 2, 3, 4 and Greek indices the values 2, 3, 4.

We know from Birkhoff's theorem (**3**), that the local field (2) can be transformed into a static Schwarzschild field. In different coordinates, the metric (2) can then be written

$$ds^2 = -e^{-\alpha}d\rho^2 - \rho^2(d\theta_1^2 + \sin^2\theta_1 d\phi_1^2) + c^2e^{\alpha}d\tau^2, \tag{3}$$

with

$$e^{\alpha} = 1 - 2\gamma m/c^2\rho - \Lambda\rho^2/3, \tag{4}$$

where m is the gravitational mass, γ the constant of gravitation, and Λ the cosmological constant.

From the boundary conditions at Σ we shall determine in Section 3 the transformation from the local coordinates $(\rho, \theta_1, \phi_1, \tau)$ to the cosmic coordinates (r, θ, ϕ, t). Let this transformation have the form

$$\rho = \rho(r, t),\ \theta_1 = \theta,\ \phi_1 = \phi,\ \tau = \tau(r, t), \tag{5}$$

and let us also assume that

$$\rho(r, t) = Rr. \tag{6}$$

Writing ρ', $\dot{\rho}$, τ', $\dot{\tau}$ for the partial derivatives of ρ, τ with respect to r, t, we then find

$$e^{\lambda} = e^{-\alpha}\rho'^2 - c^2e^{\alpha}\tau'^2,\ e^{\omega} = \rho^2,\ e^{\nu} = c^2e^{\alpha}\dot{\tau}^2 - e^{-\alpha}\dot{\rho}^2, \tag{7}$$

and

$$c^2e^{2\alpha}\tau'\dot{\tau} = \rho'\dot{\rho}. \tag{8}$$

The conditions on g_{mn} at Σ give

$$[e^{-\alpha}\rho'^2 - c^2e^{\alpha}\tau'^2] = R^2[f^{-2}], \tag{9}$$

$$[c^2e^{\alpha}\dot{\tau}^2 - e^{-\alpha}\dot{\rho}^2] = c^2, \tag{10}$$

the brackets $[\ldots]$ indicating that the quantity is evaluated at $r = a$.

The components of the stress-energy tensor for the field of (3) are zero, and hence $T^m_n = 0$ for (2). The conditions on T^m_n at $r = a$ will then be satisfied if

$T_1^1 = 0$ for the field of (1). This means that the universe is one of zero pressure and that R satisfies the equation

$$\left(\frac{2\ddot{R}}{R} + \frac{\dot{R}^2}{R^2}\right)\frac{1}{c^2} + \frac{1}{R^2R_0^2} - \Lambda = 0. \tag{11}$$

This has the integral

$$\dot{R}^2 = \frac{K}{R} - \frac{c^2}{R_0^2} + \frac{\Lambda}{3}c^2R^2, \tag{12}$$

where K is a constant of integration.

The components of T_n^m for (2) have been calculated by Sen (**10**), and are given by

$$\left.\begin{aligned}
-\kappa T_1^1 &= -e^{-\omega} + e^{-\lambda}(\tfrac{1}{2}\omega'\nu' + \tfrac{1}{4}\omega'^2) + e^{-\nu}(-\tfrac{3}{4}\dot{\omega}^2 + \tfrac{1}{2}\dot{\omega}\dot{\nu} - \ddot{\omega}) + \Lambda,\\
-\kappa T_2^2 &= -\kappa T_3^3 = e^{-\lambda}(\tfrac{1}{2}\omega'' + \tfrac{1}{4}\omega'^2 + \tfrac{1}{2}\nu'' + \tfrac{1}{4}\nu'^2 - \tfrac{1}{4}\omega'\lambda' - \tfrac{1}{4}\nu'\lambda' + \tfrac{1}{4}\omega'\nu')\\
&\quad + e^{-\nu}(-\tfrac{1}{2}\ddot{\lambda} - \tfrac{1}{4}\dot{\lambda}^2 - \tfrac{1}{2}\ddot{\omega} - \tfrac{1}{4}\dot{\omega}^2 + \tfrac{1}{4}\dot{\lambda}\dot{\nu} + \tfrac{1}{4}\dot{\omega}\dot{\nu} - \tfrac{1}{4}\dot{\lambda}\dot{\omega}) + \Lambda,\\
-\kappa T_4^4 &= -e^{-\omega} + e^{-\lambda}(\omega'' + \tfrac{3}{4}\omega'^2 - \tfrac{1}{2}\omega'\lambda') - e^{-\nu}(\tfrac{1}{2}\ddot{\lambda}\dot{\omega} + \tfrac{1}{4}\dot{\omega}^2) + \Lambda,\\
-\kappa T_4^1 &= -\dot{\omega}' - \tfrac{1}{2}\dot{\omega}\omega' + \tfrac{1}{2}\dot{\lambda}\omega' + \tfrac{1}{2}\dot{\omega}\nu',
\end{aligned}\right\} \tag{13}$$

where the Einstein constant κ has the value $8\pi\gamma/c^2$ in the units employed here. The conditions (9) and (10) are equivalent to $[e^\lambda] = R^2[f^{-2}]$, $[e^\nu] = c^2$, and (6) gives $e^\omega = R^2r^2$. Since T_1^1 is identically zero, we find from the formula in (13) that $[\nu'] = 0$. The boundary condition on $\partial g_{44}/\partial r$ is therefore satisfied. Also the boundary conditions on $\partial g_{22}/\partial r$, $\partial g_{33}/\partial r$ are satisfied on account of (6), and the other conditions on the stress-energy tensors are clearly satisfied. When the equations (6), (8) and (12) are satisfied, the boundary conditions therefore reduce to the conditions (9) and (10).

From the conditions (9) and (10)

$$[c^2e^{\lambda}\tau'\dot{\tau}]^2 = R^2[e^{-\lambda} - f^{-2}][c^2 + e^{-\lambda}\dot{R}^2a^2], \tag{14}$$

and from (8)

$$[c^2e^{\lambda}\tau'\dot{\tau}] = R\dot{R}a[e^{-\lambda}]. \tag{15}$$

Substituting (15) in (14) we find

$$[e^{\lambda}] = [f^2] - \dot{R}^2a^2c^{-2}, \tag{16}$$

and from (12) we obtain

$$[e^{\lambda}] = 1 - Ka^2/c^2R - \Lambda R^2a^2/3. \tag{17}$$

From (4), (6) and (17) we deduce

$$m = Ka^3/2\gamma. \tag{18}$$

When $K = 0$ we find from (18) that $m = 0$, unless the spherical surface bounding the hole is at an infinite distance, so that Ka^3 is finite when $K = 0$, $a = \infty$. The values of R, determined by (12), give in this case the stationary models of the universe. Accordingly it is only possible, in this case, to have a star of gravitational mass m, whose field satisfies the given boundary conditions, for the spatially open models. This includes the spatially open forms of the de Sitter universe, but not the Einstein universe.

3. *The transformation to cosmic coordinates.*—We shall now complete the determination of the transformation from local to cosmic coordinates. The transformation of ρ is given by (6), and the transformation of τ is given by the

solution of (8) which satisfies the boundary conditions (9) and (10). These conditions are not independent when equation (18) is satisfied, (10) for example being deducible from (9) and the equation (8). We find that they may be replaced by either of the conditions

$$[\tau'] = \pm R\dot{R}ac^{-2}[f^{-2}][f - f^{-1}\dot{R}^2a^2c^{-2}]^{-1}, \quad [\dot{\tau}] = \pm[f - f^{-1}\dot{R}^2a^2c^{-2}]^{-1}. \tag{19}$$

Also from (4), (6) and (8) we find

$$\tau'\dot{\tau} = c^{-2}R\dot{R}r(1 - 2\gamma m/c^2Rr - \Lambda R^2r^2/3)^{-2}. \tag{20}$$

In order to find the solution of (20) satisfying the conditions (19) it will be convenient to write τ in the form

$$\tau = V(\rho, R). \tag{21}$$

Substituting from (6) and (21) in (20) we obtain the equation

$$\left(\frac{\partial V}{\partial \rho}\right)^2 + \frac{R}{\rho}\frac{\partial V}{\partial R}\frac{\partial V}{\partial \rho} = c^{-2}e^{-2\alpha}, \tag{22}$$

and from (19) the boundary conditions at $\rho = Ra$ become

$$\frac{\partial V}{\partial \rho} = \pm \dot{R}ac^{-2}[fe^{\alpha}]^{-1}, \quad \frac{\partial V}{\partial R} = \pm[f\dot{R}]^{-1}. \tag{23}$$

Equation (22) has a solution

$$V = \int\left\{-\frac{A}{2\rho} \pm \left(\frac{A^2}{4\rho^2} + c^{-2}e^{-2\alpha}\right)^{1/2}\right\}d\rho + A\log R + B, \tag{24}$$

where A and B are constants, and the conditions (23) give

$$A = \pm R/[f]\dot{R}. \tag{25}$$

It is necessary to take the upper sign in each of the equations (19), (23), (24) and (25) in order that $d\tau$ shall be positive for positive values of dt.

To find the general integral we write $A = \chi$ in (24) and (25), and assume that $B = \phi(\chi)$. χ may now be a function of the coordinates chosen to satisfy the boundary conditions. Thus the general integral is found by eliminating χ between the equations

$$F(\tau, \rho, R, \chi) = \tau - \int\left\{-\frac{\chi}{2\rho} + \left(\frac{\chi^2}{4\rho^2} + \frac{1}{c^2}\left(1 - \frac{2\gamma m}{c^2\rho} - \frac{\Lambda\rho^2}{3}\right)^{-2}\right)^{1/2}\right\}d\rho$$
$$- \chi\log R - \phi(\chi) = 0, \tag{26}$$

$$\frac{\partial F(\tau, \rho, R, \chi)}{\partial \chi} = 0, \tag{27}$$

when the boundary condition is

$$\chi = R/[f]\dot{R}. \tag{28}$$

These equations determine the transformation of τ, when R is a solution of (12) and ρ transforms according to (6).

4. *Discussion of the solution.*—(i) When R is a solution of equation (12), the field equations give for the line element (1), the proper density of matter

$$\rho_{00} = 3K/8\pi\gamma R^2. \tag{29}$$

From (18) and (29) we find that the hole containing the star has a radius, $\rho_s = Ra$, given by

$$\rho_s = (3m/4\pi\rho_{00})^{1/3}. \tag{30}$$

This result agrees with a formula derived and discussed by Schücking (**6**) (equation (35)).

(ii) Assuming that the star has condensed from a uniform distribution of matter in the universe of density ρ_{00}, its proper mass will be

$$M = 4\pi R^3 \int_0^a \rho_{00} r^2 (1 - r^2/R_0^2)^{-1/2} dr$$

$$= m + \frac{3}{10}\left(\frac{a}{R_0}\right)^2 m + \ldots \tag{31}$$

approximately, when terms of higher order in $(a/R_0)^2$ are neglected. Only in the case $R_0 = \infty$ will there be equality between the gravitational and the rest mass of the star.

It has been shown that the exterior field of a spherically symmetric cluster of stars describing circular orbits is given by the Schwarzschild solution (**11, 12, 13**).* This field can then be joined to the field of the expanding universe as in the case of a star. Our results are therefore also applicable to the case of a globular cluster of stars, or a spherical galaxy, in the expanding universe.

In (**13**) it was shown, in the case $\Lambda = 0$, that the condition for stability in the orbits, for the stars of a cluster, led to the condition $0 < (M - m)/M < 0{\cdot}057$. Using (31) we find that these conditions are satisfied for a cluster which has condensed from a uniform distribution of matter, provided that R_0 is real and $a/R_0 < 0{\cdot}4$.

(iii) It has been shown by Pirani (**14**) that a free particle in the Schwarzschild field will spiral away from the central mass if its distance from the mass is greater than

$$\rho_m = (3\gamma m/\Lambda c^2)^{1/3}. \tag{32}$$

Whereas ρ_m is fixed, the radius ρ_s of the hole containing the star increases as the epoch increases. Hence, unless $\Lambda \leqslant 0$, the relation $\rho_s > \rho_m$ will be satisfied after a certain epoch. Any particle of the homogeneous distribution suffering a perturbation such that it enters the hole, will then describe a path which will eventually take it further from the star. Thus the star cannot accrete matter from the homogeneous distribution. From (30) and (32) the condition for this is

$$\Lambda > 4\pi\gamma\rho_{00}/c^2. \tag{33}$$

When (33) is satisfied, the equations (11), (12) and (29) give $\ddot{R}/R > 0$ and therefore the rate of expansion of the universe is increasing.

(iv) The Einstein–Straus boundary conditions require g_{mn}, $\partial g_{mn}/\partial r$ to be continuous at $r = a$ for all t. That these conditions are not satisfied by the solution obtained above can be shown in the following manner. All the Einstein–Straus conditions are included in the O'Brien–Synge conditions except the continuity of $\partial g_{11}/\partial r$ at $r = a$. This condition leads to

$$[\lambda'] = 2a[f^{-2}]R_0^{-2}. \tag{34}$$

The value of $[\lambda']$, for the solution obtained above, can be found from the equation $T_4^4 = 0$, for the line element (2), at $r = a$. Substituting the values of the functions and their derivatives at $r = a$ in the expression for T_4^4 given in (13), it is found that (34) can only be satisfied when $K = 0$.

* The only difference the O'Brien-Synge boundary conditions make to the solution in (**13**), is that the condition on ρ_{00} following equation (19) is no longer necessary.

It should, however, be noted that Einstein and Straus used isotropic coordinates in their investigation and the solution which is obtained in those coordinates by use of the Einstein–Straus boundary conditions will necessarily satisfy the O'Brien–Synge conditions. Such a solution will also satisfy the Einstein–Straus conditions in the non-isotropic coordinates of the metric (1). The different solution which satisfies the O'Brien–Synge conditions in non-isotropic coordinates is one for which the line elements (1) and (2) cannot both be transformed to the isotropic form by a single transformation.

The fact that Schücking (**6**) has used the transformation (6) and that he also obtains a formula agreeing with (30) makes it probable that his solution satisfies the O'Brien–Synge conditions and not the Einstein–Straus conditions.

Department of Mathematics,
King's College,
Newcastle-upon-Tyne:
1956 *October* 7.

References

(1) A. Einstein and E. G. Straus, *Rev. Mod. Phy.*, **17,** 120, 1945.
(2) A. Einstein and E. G. Straus, *Rev. Mod. Phy.*, **18,** 148, 1946.
(3) G. D. Birkhoff, *Relativity and Modern Physics*, 253–256, 1923.
(4) B. Qvist, *Soc. Sci. fenn.*, Comm. Phys.-Math., **13,** No. 11, 1947.
(5) P. Kustaanheimo, *Soc. Sci. fenn.*, Comm. Phys.-Math., **13,** No. 12, 1947.
(6) E. Schücking, *Zs. f. Physik*, **137,** 595, 1954.
(7) S. O'Brien and J. L. Synge, Comm. of the Dublin Inst. for Advanced Studies, A., No. 9, 1952.
(8) W. B. Bonnor, *Zs. f. Astrophysik*, **35,** 10, 1954.
(9) W. B. Bonnor, *Zs. f. Astrophysik*, **39,** 143, 1956.
(10) N. R. Sen, *Zs. f. Astrophysik*, **10,** 291, 1935.
(11) A. Einstein, *Annals of Math.*, **40,** 922, 1939.
(12) N. R. Sen and T. C. Roy, *Zs. f. Astrophysik*, **34,** 84, 1954.
(13) C. Gilbert, *M.N.*, **114,** 628, 1954.
(14) F. A. E. Pirani, *Proc. Camb. Phil. Soc.*, **50,** 637, 1954.

DIRAC'S COSMOLOGY AND THE GENERAL THEORY OF RELATIVITY

C. Gilbert

(Received 1956 November 9)

Summary

An explanation is given, according to the principles of the general theory of relativity, of some results previously obtained by Dirac. It is assumed that there is a different unit of distance for electromagnetic phenomena from the unit of distance used for describing gravitational phenomena. The ratio of these units varies with the time. The former unit gives the usual "cosmic" distance and a Newtonian law of gravitation which is independent of the time, the latter unit leads to a Newtonian law of gravitation for which the gravitational power of matter varies inversely as the epoch.

Formulae are derived which enable the age of the universe ($\bar{t}$), Hubble's constant (H) and the mean density of matter in the universe (ω), to be calculated from the values of the constants of atomic theory and the present observed value of the gravitational "constant". It is found that $\bar{t}=4{\cdot}1\times10^{9}$ years, $H=160$ km/sec/megaparsec, and $\omega=4{\cdot}8\times10^{-29}$ gm/cm^3.

---

1. *Introduction.*—Nearly twenty years ago Dirac gave "a new basis for cosmology" (**1**), in which he used some numerical coincidences which occur between large dimensionless numbers formed from the constants of Nature, to show that the gravitational "constant" is in fact not constant, but a quantity γ which varies inversely as the epoch. Dirac gave a relativistic treatment of the theory, but unfortunately the method he used is open to criticism (**2**), and consequently his results have not been generally accepted. It is, in fact, generally believed that a theory of gravitation with variable γ is beyond the scope of the general theory of relativity. In this paper I show that this is not the case and I obtain results similar to those given by Dirac by use of the principles of the general theory of relativity.

The usual derivation of the Einstein constant κ depends on obtaining an approximation to Poisson's equation for weak static fields, without considering the relation of these fields to the field generated by the rest of the matter in the universe. This has led to the assumption having been made that the spatial coordinates and the time occurring in Newtonian theory are the same as those for which the velocity of light *in vacuo* is constant. It will be shown that this is not necessarily the case, and, in order to bring this point out clearly, we derive the value of κ in a different manner, which makes use of equations for the relativistic models of the universe, which form the basis of "Newtonian Cosmology" (**3**, **4**). We then show how Newton's law of gravitation can be derived from the relativistic equations of motion, for both local and large-scale gravitational phenomena, in two forms. In one form the usual "cosmic" coordinates are used and the gravitational power of matter is constant, in the other form a different unit of distance is used and the gravitational power γ varies inversely as the epoch.

Classical electrodynamics is independent of Newtonian mechanics. It is not therefore necessary for the spatial coordinates used in forming Maxwell's equations *in vacuo* to be the same as those occurring in Newton's law of gravitation. We shall assume that the former coordinates correspond to *cosmic* distance coordinates and the latter coordinates are those which give rise to the value of γ which varies inversely as the epoch. We also assume that the measurements of the coordinates and other quantities occurring in the description of the Universe are based on atomic units of mass, length, and time. The rather tentative nature of the approach, which comes from making these assumptions, is made more rigorous in Section 10, where it is shown that Dirac's equation expresses the relation between the units of the electrical and gravitational fields in a local gauge-system.

We consider the Einstein–de Sitter model of the universe (**5**), and we show that the present epoch may be calculated for this model, from the present values of the physical constants. The properties of the model may then be fully determined without making assumptions about the values of any arbitrary constants.

P. Jordan (**6, 7**) and E. A. Milne (**8**), have described models for which the physical constants varied with the time, but the present approach is different from theirs, which used different mathematical techniques. The cosmology of Jordan is based on Dirac's results, and uses the technique of projective relativity, whereas the rather different results of Milne's theory were obtained by the methods of kinematic relativity.

2. *The atomic units of measurement.*—We assume that the metric of the expanding universe has the form

$$ds^2 = c^2\,dt^2 - R^2(dx_1^2 + dx_2^2 + dx_3^2), \tag{1}$$

where R is a function of t only, c is the velocity of light and the 3-spaces $t =$ constant are conformal to Euclidean space in which radial distance $r = (x_1^2 + x_2^2 + x_3^2)^{1/2}$.

We assume that the matter of the universe exerts negligible pressure and that the cosmological constant occurring in the field equations is zero. We then find from the field equations that the mean density of matter ρ_{00} is given by

$$\kappa\rho_{00} = 3\dot{R}^2/R^2c^2, \tag{2}$$

where κ is Einstein's constant, and

$$2\frac{\ddot{R}}{R} + \frac{\dot{R}^2}{R^2} = 0. \tag{3}$$

From (3) we find

$$\dot{R}^2 = K/R, \tag{4}$$

where K is a constant of integration. A further integration gives

$$R = (\tfrac{3}{2}K^{1/2}t + A)^{2/3}, \tag{5}$$

where A is a constant. We can without loss of generality take $A = 0$ and

$$K = \tfrac{4}{9}C^2. \tag{6}$$

Equation (5) then gives

$$R = (Ct)^{2/3}. \tag{7}$$

The system of coordinates (t, x_1, x_2, x_3) we call *basic* coordinates. These coordinates are usually said to be co-moving because the coordinate distance between any two particles of matter is constant. We assume that although the

matter in the universe may be treated mathematically as a continuous distribution, it does in fact consist of protons and electrons of masses m_p and m_e gm and that associated with these particles there are fundamental intervals of time ϵ secs and distance δ cm. The basic coordinates are assumed to be measured in terms of these fundamental constants as units of time and length. We also assume that the average charge density in the neighbourhood of any event is zero, and that t at any event is measurable from the local properties of matter, and therefore has the nature of an absolute time and gives a meaning to simultaneity of events independently of light-signals between events.

3. *The value of Einstein's constant.*—The equation governing the radial motion of matter in the universe can be put in Newtonian form when the distance coordinates are

$$X_\alpha = Rx_\alpha \quad (\alpha = 1, 2, 3). \tag{8}$$

Writing

$$\rho = (X_1^2 + X_2^2 + X_3^2)^{1/2} = (Ct)^{2/3} r, \tag{9}$$

and $v = d\rho/dt$, we find that

$$v = \frac{2}{3}\frac{\rho}{t}. \tag{10}$$

The equation giving the radial motion of matter is then

$$\frac{dv}{dt} = -\frac{2}{9}\frac{\rho}{t^2}. \tag{11}$$

Writing $M(\rho)$ for the mass contained within a sphere of radius ρ, we find from (2), (4), (6) and (7)

$$M(\rho) = \frac{16\pi}{9\kappa c^3}\frac{\rho^3}{t^2}, \tag{12}$$

$$\frac{dv}{dt} = -\frac{\Gamma M(\rho)}{\rho^2}, \tag{13}$$

where Γ is a constant satisfying

$$\kappa = \frac{8\pi\Gamma}{c^2}. \tag{14}$$

An observer who uses coordinates (t, X_1, X_2, X_3) and uses Euclidean space will describe the system as subject to a Newtonian law of gravitation, with gravitational constant Γ, provided he assumes that matter at distances greater than ρ gives no resultant gravitational force on the matter inside a sphere of radius ρ. This result, which was obtained by Milne (3), is used in the present instance to obtain the relation (14) defining κ.

4. *The local gravitational field.*—The coordinates (t, X_1, X_2, X_3) are usually called *cosmic* coordinates. It has been shown (9) that when *cosmic* spatial coordinates are used the gravitational field of a particle of mass m in the universe is of the static Schwarzschild form, having the line element

$$ds^2 = c^2 e^\nu d\tau^2 - e^{-\nu} d\rho^2 - \rho^2(d\theta^2 + \sin^2\theta\, d\phi^2), \tag{15}$$

where

$$e^\nu = 1 - \frac{2\Gamma m}{c^2\rho}, \tag{16}$$

and (ρ, θ, ϕ) are the spherical polar coordinates corresponding to cosmic coordinates (X_1, X_2, X_3) for Euclidean space.

The condition that the field (15) shall join continuously onto the field (1) at the 3-space $r = a$ has been shown to be (**9**)

$$m = \frac{Ka^3}{2\Gamma}. \tag{17}$$

We assume that Γ is chosen to make $a = 1$ when $m = 1$. From (6) and (17) we then find

$$\Gamma = \frac{2}{9}C^2. \tag{18}$$

The time τ in (15) in general differs from cosmic time, but for speeds very much less than the speed of light, and when $\Gamma m/c^2\rho \ll 1$, $d\tau$ is approximately equal to ds/c, which is the cosmic time measured by an observer at himself. It is also known that for speeds very much less than the speed of light the equations of motion for a free particle in the field (15) have the Newtonian form when ds/c plays the part of Newtonian time. *Cosmic* time is therefore approximately the same as Newtonian time for local gravitational phenomena.

5. *The velocity of light.*—The equations of the general theory of relativity governing both the large scale motions of matter in the universe and the orbital motions about a condensation of matter, have been shown to agree with the equations of motion derived from Newtonian theory, when *cosmic* coordinates are used. We now show that the relativistic equations giving the propagation of light *in vacuo* are in agreement with Maxwell's equations provided that *cosmic* time and *cosmic* spatial coordinates are used in the latter equations.

Consider a spherical wave emitted from $r = 0$ at time $t = t_0$. Since the light paths are the null geodesics of (1), we find that in basic coordinates the wave front at time t is given by

$$r = 3cC^{-2/3}(t^{1/3} - t_0^{1/3}). \tag{19}$$

From (7) and (19) the wave front in *cosmic* coordinates is

$$\rho = 3c(t - t^{2/3}t_0^{1/3}). \tag{20}$$

From (20) we find that the velocity of light in cosmic coordinates tends to the limiting value $3c$ as t tends to infinity. When $t = \infty$, the coordinate density of matter is zero, and therefore when *cosmic* coordinates are used the velocity of light *in vacuo* is $3c$.

The constant C in (7) was arbitrary. We now assume that it is a dimensionless number formed from the constants of atomic theory. The cosmic coordinates are then expressed in terms of constants of atomic theory. Since it is always possible to form a quantity having the dimensions of a length from one having the dimensions of time by multiplication by the velocity of light, it is possible to choose one of the quantities ϵ and δ so that the velocity of light in cosmic coordinates is unity. Hence we can take

$$c = \tfrac{1}{3}. \tag{21}$$

6. *The variation of the gravitational power of matter with time.*—We now show that the relativistic equations can be put into the Newtonian form when spatial coordinates differing from *cosmic* coordinates are used. In these coordinates γ is found to vary with the time. Let these coordinates ζ_α ($\alpha = 1, 2, 3$) be defined by

$$\zeta_\alpha = \lambda x_\alpha, \tag{22}$$

where λ is a slowly varying function of t, and write $\zeta=(\zeta_1^2+\zeta_2^2+\zeta_3^2)^{\frac{1}{2}}$, $\bar{v}=d\zeta/dt$. We then find that the radial motion of the matter in the universe is given by

$$\frac{d\bar{v}}{dt}=\frac{\ddot{\lambda}}{\lambda}\zeta. \tag{23}$$

Writing $M(\zeta)$ for the mass of the homogeneous distribution contained within a sphere of radius ζ we find from (12) and (22) that

$$M(\zeta)=\frac{\zeta^3}{\lambda^3} \tag{24}$$

and

$$\frac{d\bar{v}}{dt}=-\frac{\gamma M(\zeta)}{\zeta^2}, \tag{25}$$

where

$$\gamma=-\lambda^2\ddot{\lambda}. \tag{26}$$

Since λ varies slowly with the time we find for the local field (15) that a free particle will move under a Newtonian law of gravitation, approximately, with

$$\gamma=\frac{\lambda^3}{R^3}\Gamma=\frac{2}{9}\frac{\lambda^3}{t^2}. \tag{27}$$

Equating the values of γ from (26) and (27) we find that λ must satisfy the equation

$$\frac{\ddot{\lambda}}{\lambda}=-\frac{2}{9t^2}. \tag{28}$$

Integration of (28) gives

$$\lambda=C_1t^{2/3}+C_2t^{1/3}, \tag{29}$$

where C_1 and C_2 are arbitrary constants.

7. The case $C_1=C^{2/3}$, $C_2=0$ gives the *cosmic* system of coordinates, which have already been discussed. We now discuss the case $C_1=0$, $C_2=1$ giving coordinates $(t, \zeta_1, \zeta_2, \zeta_3)$, which we shall call *local* coordinates.

We have from (22)

$$\zeta_\alpha=t^{1/3}x_\alpha, \tag{30}$$

and from (27)

$$\gamma=\frac{2}{9}\cdot\frac{1}{t}. \tag{31}$$

This shows that in *local* coordinates the gravitational power of matter varies inversely as the epoch, in agreement with a result obtained by Dirac (1).

8. We assume that a terrestial observer uses *cosmic* coordinates for the description of electromagnetic phenomena and *local* coordinates for the description of gravitional phenomena. Also instead of using the atomic units of length mass and time he uses the c.g.s. system of measurement in which the atomic units have the values δ, m_p and ϵ respectively.

Writing $\bar{c}$, $\bar{t}$ and $\bar{\gamma}$ for the velocity of light, the epoch and the gravitational power of matter in the c.g.s. system we have

$$\bar{c}=\frac{\delta c}{\epsilon}=\frac{\delta}{3\epsilon}, \tag{32}$$

$$\bar{t}=\epsilon t, \tag{33}$$

$$\bar{\gamma}=\frac{\delta^3}{m_p\epsilon^2}\gamma, \tag{34}$$

From (31), (32), (33) and (34) we find

$$\bar{\gamma} = \frac{6\epsilon^2 \bar{c}^3}{m_p \bar{t}}. \tag{35}$$

The quantity ϵ has to be chosen from amongst the quantities having the dimensions of time which can be formed from the constants of atomic theory. We take $\epsilon = e^2/m_e\bar{c}^3$, where e, m_e are the charge and mass of an electron. This was the unit of time taken by Dirac and we find that it leads to numerical results which are in good agreement with observation. Substituting this value for ϵ in (35) we find

$$t = \frac{\bar{t}}{\epsilon} = \frac{6e^2}{m_p m_e \bar{\gamma}}. \tag{36}$$

9. *Comparison with observation.*—From the present value of the gravitational "constant" and the atomic constants, we can calculate the value of $\bar{t}$ from (36). We call this value of $\bar{t}$ the "age of the universe". We find that

$$\textit{the age of the universe } \bar{t} = 4{\cdot}1 \times 10^9 \text{ years}. \tag{37}$$

It can be shown that *cosmic* distance corresponds closely to "luminosity distance" (2) and must therefore be used in making estimates of Hubble's constant and the mean density of matter in the universe.

From (10) we find that Hubble's constant is given by

$$H = \frac{2}{3\bar{t}}. \tag{38}$$

Substituting for $\bar{t}$ from (36) we find that

$$\textit{Hubble's constant } H = 160 \text{ km/sec/megaparsec}, \tag{39}$$

corresponding to an age calculated from its reciprocal of 6×10^9 years.

The mean density of matter in the universe, ω, is the same as the proper density of matter, when cosmic distance is used. Therefore

$$\omega = \frac{1}{6\pi\bar{\gamma}\bar{t}^2} \text{ protons/unit cosmic volume} \tag{40}$$

$$= \frac{1}{6\pi\bar{\gamma}\bar{t}^2} \text{ gm/cm}^3. \tag{41}$$

This gives

$$\textit{the mean density of matter } \omega = 4{\cdot}8 \times 10^{-29} \text{ gm/cm}^3. \tag{42}$$

10. *The local gauge-system.*—The following generalization may be regarded as the essentially new concept of the work which has been done. *The measurements of all quantities occurring in field theory shall be based on units of mass, length and time, derived from the values of the atomic "constants" measured in a local frame with arbitrarily chosen units.* In order to discuss this generalization with regard to the Einstein–de Sitter universe we first make a transformation $t = C^2T^3$, where T is a new basic time coordinate. We find from (1), (7) and (21),

$$ds^2 = (CT)^4(dT^2 - dx_1^2 - dx_2^2 - dx_3^2). \tag{43}$$

Let the proper interval in a local frame with arbitrarily chosen units be given by

$$ds'^2 = \bar{c}^2 dT'^2 - dx_1'^2 - dx_2'^2 - dx_3'^2, \tag{44}$$

and let the measured values of the atomic constants be $\bar{c}$, e, m_e, m_p. If we choose new units of mass, length and time m_p, $e^2/m_e\bar{c}^2$ and $e^2/m_e\bar{c}^3$ the measured values

of the atomic constants are 1, $\pm(m_e/m_\mu)^{1/2}$, m_e/m_μ, 1, respectively, and in basic coordinates (44) becomes

$$ds_0^2 = dT^2 - dx_1^2 - dx_2^2 - dx_3^2 \tag{45}$$

where $ds' = (e^2/m_e\bar{c}^2)\,ds_0$. The atomic constants are therefore absolute constants for basic coordinates provided m_e/m_μ is an absolute constant, and their values will not be affected by transformations of gauge. The electric field of a stationary charge will thus transform differently from the gravitational field of a particle under a gauge transformation. From the point of view of field theory Dirac's equation determines the gauge-system in which the unit of the electrical field is defined in terms of the unit of the gravitational field. In the preceding work measurements of distances in this gauge-system have been made in *local* coordinates, and the corresponding interval of proper time in the local gauge is $ds_1 = ds/C^{4/3}T$. It is possible that this local gauge-system may be determined by a condition for the existence of free charges of the type proposed by Dirac (**10**).

In the uncharged state of the universe which we have been discussing, gauge transformations are merely transformations to conformal space-times which do not alter the field equations. The structure of the universe is then the same in the different gauge-systems, but its description is different. It is from the description of the local gravitational field that we have been able to determine the local gauge-system without a full knowledge of the field theory.

The generalization which has been arrived at is derived from the necessity for field theory to use a system of measurement in which the atomic constants are absolute constants. The suggestion of M. Born (**11**, **12**) that field theory should "introduce an absolute length $(e^2/m_e\bar{c}^2)$ right from the beginning" is probably closely related, but was made for different reasons.

Department of Mathematics,
King's College,
Newcastle-upon-Tyne:
1956 November.

References

(1) P. A. M. Dirac, *Proc. Roy. Soc.* A, **165**, 199, 1938.
(2) H. Bondi, *Cosmology*, 162, 108, C.U.P., 1950.
(3) E. A. Milne, *Quart. Journ. of Math.*, **5**, 64, 1934.
(4) W. H. McCrea and E. A. Milne, *Quart. Journ. of Math.*, **5**, 73, 1934.
(5) A. Einstein and W. de Sitter, *Proc. Nat. Acad. of Sciences*, **18**, 213, 1932.
(6) P. Jordan, *Die Herkunft der Sterne*, Stuttgart, 1947.
(7) P. Jordan, *Nature*, **164**, 637, 1949.
(8) E. A. Milne, *Kinematic Relativity*, Oxford, 1948.
(9) C. Gilbert, *M.N.*, **116**, 678, 1956.
(10) P. A. M. Dirac, *Proc. Roy. Soc.* A, **209**, 291, 1951.
(11) M. Born, *Nature*, **169**, 1105, 1952.
(12) M. Born, *Rev. of Mod. Phys.*, **21**, 463, 1949.

A NEW EXPLANATION OF THE RECESSION OF THE NEBULAE

E. R. R. Holmberg

(Received 1956 October 5)*

Summary

The basis of the explanation is the proposal that the space appropriate to cosmology has three, not four, dimensions. The coordinates x, y, z, ct of the so-called light cone are taken to be homogeneous coordinates of a projective space which is then specified to have a uniform elliptic metric. This involves replacing the time coordinate by

$$W = \sqrt{R^2 - c^2t^2}$$

in a manner analogous to the replacement of classical kinetic energy by the relativistic quantity

$$mc^2/\sqrt{1 - v^2/c^2}.$$

The radius of the universe, R, is thus introduced explicitly. The components of the 4-velocity vector are shown to be interpretable as the plane coordinates of this space.

Clifford's work on the parallels of elliptic space is recalled and used to show that the physical identifications above of a point and plane manifold in elliptic space would lead one to expect the illusion of a recession velocity proportional to distance.

---

1. *Introduction.*—Although cosmology provides the simplest field of application for the theory of relativity—the metric is uniform—no completely satisfactory explanation of the observed facts in terms of it has yet been put forward.

Three stages in the development of ideas can be traced. In the first stage the metric is supposed to be isotropic with respect to all four dimensions of space and time. No solution of this type can be found unless an arbitrary negative pressure is introduced to stop the universe expanding. In the second stage the isotropy condition with respect to the time coordinate is dropped and the universe is allowed to expand. Support for this solution is found in the Hubble red shift of the spectral lines in the light from distant stars. In order to avoid the implications of a starting point in the expansion of the universe the third stage is to assume a steady state in which the loss of matter from any one neighbourhood caused by the expansion is made good by continual creation.

Since the discovery of the Hubble red shift and the analogous effect in the 21 cm line of radio astronomy, both indicating proportionality of recession velocity and distance, the static solution of the cosmological equations has had to be abandoned. Unfortunately neither the simple expanding universe hypothesis nor its continuous creation variant are without their difficulties. The latter in particular has recently been threatened by the radio-astronomical observations of Ryle (1) which seem to indicate that at very great distances there is a crowding together of matter more in accord with the simple hypothesis of expansion: this crowding together can be taken as evidence for a change in size

* Received in original form 1956 July 12.

of the universe with time, because the distant stars are seen at an epoch in the past, whereas the continuous creation hypothesis apparently implies complete isotropy to the limit of the assessible universe. On the other hand the simple expansion hypothesis fails to account for the fact that in every neighbourhood of space stars can be found at all stages of development: if the universe is not in a steady state one would expect to find the younger stars farther away, that is, back in the past.

Besides their disagreement with some of the observational evidence neither of the expanding universe hypotheses finds easy acceptance on intuitive and philosophical grounds. Both imply creation of one kind or another and although continuous creation can be represented as a process ultimately capable of mathematical description it is difficult at the moment to see how it will be reconciled with the established laws of physics. The need for such drastic proposals would be avoided however if it could be shown that the effects like the expansion of the universe and the crowding together of distant matter were only illusions inherent in the methods of observation allowed us. For if these effects do not correspond to actually occurring physical processes we must not make the usual physical deductions from them. We could then revert to a static or stable cosmological solution in which the universe has no beginning and no end but, being closed and isolated and therefore losing no energy, is in a state of perpetual change with stationary characteristics.

The possibility of finding such a solution, and of showing the illusory nature of some of the astronomical evidence, is offered by the circumstance that all the cosmological theories mentioned have in them a certain arbitrary feature. The negative pressure of the original theory, the ratio of the curvature of space to the extent of the accessible neighbourhood set by the velocity of light and the rate of expansion of the universe in the later ones. The existence of these arbitrary constants indicates that the basic theory on which cosmology is at present founded may allow for more degrees of freedom than are physically necessary or possible. Accordingly the desired result may be achieved only by a revision of the underlying theory and not by finding yet another model based on that theory as it stands at present.

2. *Metrical geometry.*—The mathematical basis for all the previous cosmologies is the metrical geometry of 4-dimensions. There are a number of ways of developing this geometry which have formerly been regarded as equivalent. From what follows in this paper, however, it seems possible that one of these ways, namely, through projective geometry and the invariant theory of Cayley (**2**), may be superior, at least insofar as it enables one to keep the status of the various axioms and definitions clearly in mind. After following this development one is led to a somewhat novel representation of the geometry of the universe which appears to remove the arbitrary elements of the earlier cosmologies described above and yet leave room for a natural explanation of most of the observational facts.

For an account of the principles of projective geometry the reader is referred to the standard textbooks,* where it will be found that the coordinates required correspond only to labels of the points, planes and other manifolds of space and that differences between pairs of coordinate sets have no invariant significance and can therefore not be considered as representing distances. The Cayley

* e.g. Baker, H. F., *Principles of Geometry*, Vol. I.

theory (**2**) of distance between two elements of a manifold defines it as the logarithm of the cross ratio of the coordinates of two elements and the coordinates of the two intersections of their join with an absolute quadric surface. This expression is invariant under the group of transformations which takes the absolute quadric into itself, and Lie (**3**) has shown that there are three such groups only corresponding to different forms of quadric. If the quadric is real and closed one obtains the hyperbolic geometry of Lobatchewsky; if it is imaginary, the elliptic geometry of Riemann; if it degenerates into a single plane the geometry is affine, and if in this plane a closed conic is latent the geometry is Euclidian.

We give now the principal result of the Cayley theory which is the basis of the proposal in this paper. It is that the distance, D, between two points x_i' and $x_i (i = 1 \ldots n+1)$ in an n dimensional manifold can be put in the form:

$$\cos(D/k) = \frac{a_{11}x_1'x_1 + a_{12}(x_1'x_2 + x_2'x_1) + a_{22}x_2'x_2 + \ldots}{\sqrt{(a_{11}x_1'^2 + 2a_{12}x_1'x_2' + \ldots)}\sqrt{(a_{11}x_1^2 + 2a_{12}x_1x_2 + \ldots)}} \qquad (2.1)$$

where

$$\phi = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + \ldots = 0$$

is the equation of the absolute quadric in homogeneous coordinates and k is an arbitrary scale factor which connects the numbers used for the coordinates and those which measure distance.

In Euclidian geometry the limiting procedures which define it and which correspond to the degeneration of the metric quadric result in the cancellation of the expressions in the denominator of (2.1) and the reduction of the numerator to a form which involves coordinate differences only. But in the non-Euclidian cases it appears impossible to reduce the R.H.S. of 2.1 to a form involving only differences of coordinate values, even when these differences are small, without explicitly setting the expressions under the root signs in the denominator equal to a constant, that is imposing the condition:

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + \ldots = k^2 \qquad (2.2)$$

on all the coordinate sets. (Incidentally, it is necessary to make this new constant equal to the scale factor on the left of 2.1 in order to preserve the usual definition of distance.) This condition can then be used to eliminate one of the coordinates, giving an expression for distance involving the same number of variables as dimensions. This reduction to inhomogeneous coordinates occurs automatically in Euclidian geometry described relative to orthogonal axes, and only in this case are the coefficients in the distance expression constants. In a sense, therefore, the passage from Euclidian to non-Euclidian geometry can be regarded as a generalization: the replacement of constant by non-constant coefficients. The cosmological theories which have developed from Einstein's relativity have all taken as their starting point a distance formula in which the reduced number of variables, that is, the number of inhomogeneous coordinates, was four, and in which the coefficients were not all constant. All these theories are thus based on the premise that the universe to be described is four-dimensional.

This paper adopts another premise, that the universe, as accessible to observation, is only three dimensional. This is not an arbitrary procedure, as will be shown. It takes its justification from the fact that the equations of special relativity can be interpreted as referring to a three-dimensional quite

as well as to a four-dimensional world. Other more cogent arguments in favour of the new premise can be adduced but these would be outside the scope of this paper. Here therefore after demonstrating that the premise need not contradict established facts we rest inductively on its logical consequences for its acceptance.

3. *A modification of special relativity.*—It will be observed immediately that the relation (2.2) for a 3-dimensional space is identical with the condition that a 4-vector in Euclidian space should have a constant length. The fact that in such a condition the principal coefficients are all unity and the cross products vanish is merely a consequence of choosing the coordinate system so that the equation of the metric quadric has the canonical form:

$$x_1^2 + x_2^2 + \ldots = 0.$$

There is therefore no formal difference between the properties of constant vectors in 4-dimensional Euclidian space and general vectors in 3-dimensional non-Euclidian space of constant curvature, and the question poses itself: whether we have been physically correct in relativity in supposing that constant 4-vectors do in fact refer to a 4-dimensional space. As remarked above, this paper is hardly the place to discuss the geometrical implications of the problem, so we shall proceed by taking the newly exposed alternative as a working hypothesis and show what its consequences are.

The specific 4-vector which concerns us is the 4-velocity, which we can write for our present purpose in its usual form:

$$v_1 = \{i\xi v/c/\sqrt{1 - v^2/c^2},\ i\eta v/c/\sqrt{1 - v^2/c^2},\ i\zeta v/c/\sqrt{1 - v^2/c^2},\ 1/\sqrt{1 - v^2/c^2}\}$$

where $\xi^2 + \eta^2 + \zeta^2 = 1$. It therefore satisfies the condition

$$v_1^2 = 1.$$

Our proposal is that the v_1 should be regarded as the homogeneous plane co-ordinates of a uniformly curved 3-space. There is no contradiction of the accepted principles of special relativity involved in doing this for the rules of composition of velocities in relativity correspond exactly to the Cayley formula (2.1). Taking as an example two velocities in the same direction (ξ, η, ζ), (2.1) becomes:

$$\cos(D/c) = \frac{1 - vv'/c^2}{\sqrt{(1 - v^2/c^2)}\sqrt{(1 - v'^2/c^2)}}$$

or

$$\tan(D/c) = \frac{i(v' - v)/c}{1 - vv'/c^2}.$$

The novel consequences of the proposal appear only because of the fact that if the velocity manifold is taken as a plane manifold it must have a point manifold associated with it, which also has non-Euclidian metrical characteristics. This is the nub of the matter and seems easily to be overlooked if any of the alternative approaches to relativity are employed. This point manifold must, like the plane manifold, be isotropic and so its sets of four homogeneous coordinates must be connected by a relation of the type (2.2). It follows therefore that we cannot be dealing here with the ordinary space-time of relativity because the coordinates of this are in principle independent.

The solution offered by the author is that the point manifold corresponding to the velocity manifold is the so-called light cone. The coordinates, x, y, z, ct of this space are not independent and can with modification be regarded as

referring to a non-Euclidian space of 3-dimensions. The initial difficulty in the way of making this identification is occasioned by the fact that as they stand x, y, z, ct satisfy

$$x^2 + y^2 + z^2 - c^2t^2 = 0 \tag{3.1}$$

in which the characteristic constant is zero. This is clearly not compatible with the Cayley theory for it would force the distance between two points to be always infinite. To overcome this one can adopt the device used by Einstein in constructing the 4-velocity, that is, to add to the time term a constant so large that it is normally neglected. By putting:

$$c^2t^2 = R^2 - w^2 \tag{3.2}$$

(3.1) becomes:

$$x^2 + y^2 + z^2 + w^2 = R^2. \tag{3.3}$$

If R is very large compared with x, y, z, ct on the Earth we find ourselves dealing with the coordinates

$$\frac{x}{R}, \quad \frac{y}{R}, \quad \frac{z}{R}, \quad 1$$

in most situations, where R can be absorbed into the scale factor and space appears to be straight. The law of compounding the coordinates also reduces to the classical form in the limit. For example (3.3) is satisfied if the coordinates are put in the form

$$\xi ct, \quad \eta ct, \quad \zeta ct, \quad \sqrt{R^2 - c^2t^2}$$

where $\xi^2 + \eta^2 + \zeta^2 = 1$. The distance between two points in the same direction is then given by

$$\cos(D/R) = \frac{1}{R^2}\{c^2tt' + \sqrt{(R^2 - c^2t^2)(R^2 - c^2t'^2)}\}.$$

When ct and ct' are small the R.H.S. reduces to:

$$1 - \frac{1}{2}\frac{c^2}{R^2}(t' - t)^2$$

so that $D = c(t' - t)$. If the sign of the spatial components of one of the sets of coordinates is made negative the same procedure yields the sum of two times instead of the difference, as it should. The sign of the time-like component must not be changed since this is in the nature of a cosine.

It is clear from the work just done that R plays an analogous role in the point manifold to the velocity of light in the plane manifold. With this difference: the plane manifold, on account of the spatial components being counted imaginary, is hyperbolic, whereas the point manifold is elliptic. In the plane manifold, therefore, the metric quadric is real and constitutes a boundary which is the velocity of light. The point manifold is, on the other hand, unbounded although finite. Its extent is a matter of choice and is settled by the scale factor R. This may be regarded as "the radius of the universe". This is not a trivial identification for it can be seen to correspond closely with what we know of the physical universe. Our experience of events outside the solar system and particularly outside our own galaxy is, because of the great distances and finite velocity of light, effectively confined to one observation at one instant of time. The time taken for light to reach us with information about these events is so great compared even with the whole of written human history that these events can be considered

as lying on a light cone. The point manifold whose geometry we have just analysed is therefore entirely appropriate for the representation of our knowledge of the physical universe and the radius of this manifold can be identified with the radius of the actual universe.

4. *A new interpretation of nebular recession.*—The properties of 3-dimensional non-Euclidian space were examined by W. K. Clifford (4) who discovered a number of significant differences between them and those of Euclidian space. At the time Clifford wrote it was customary to regard elliptic and hyperbolic spaces as distinct alternatives and, in speaking of point and plane manifolds, therefore, he regarded it possible to have both in either an elliptic or hyperbolic space. From what has been said in the last section however it is clear that the difference between elliptic and hyperbolic spaces is a matter of whether one counts all the coordinates as real or whether one (or three) of them is taken as imaginary. The fact that Clifford's results referred to an elliptic space therefore need not prevent us from applying them to the association of an elliptic point manifold and hyperbolic plane manifold we have discussed above. In fact the point of view implied by this appears the more natural one. In Newtonian space the point manifold is Euclidian and open and the plane manifold is closed. It accords with our intuitive view of things to have point and plane manifolds of a space complementary in this respect.

To return to the main argument, Clifford pointed out that in elliptic space points and planes stand in a definite relation to each other which has no parallel in Euclidian space. Thus each point defines a unique plane, its polar with respect to the metric quadric, and each plane defines a point, its pole with respect to the same quadric. It follows from this that the points of a line define a pencil of polar planes, so that each line has a polar line with respect to the metric quadric. Accordingly, translation along a line is equivalent to rotation about the polar line as axis, and rotation about the same line is equivalent to translation along the polar. Clifford also showed that two self-polar lines lie in the surface generated by translating one of them parallel to itself. There are in general two lines parallel to a given line and through a point not on it, but these coincide when the point lies on the polar line; this line is the polar itself.

From these results of Clifford's certain peculiarities in the process of interpreting events occurring at great distances from an observer seem to be indicated. For example, when translation is being observed at a distance this has to be described, as far as magnitude and direction are concerned, by referring to a set of parallel axes close at hand, so the peculiarities of parallels in elliptic space are significant. The matter is simplest to visualise in the extreme case when these parallel axes are the polars of the ones along which the actual translation takes place. We must then take into account that translation along these is equivalent to rotation about the latter. It seems plausible to suppose that this equivalence means that translation taking place at the polar distance will appear and must be represented as rotation about the local parallel axis and vice-versa. Also, at intermediate distances a translation or a rotation would need to be represented as a twist about a screw. These remarks have to be interpreted somewhat before being applied to cosmology. Clifford had in mind a simple generalization of Newtonian space as his model of the point manifold, whereas we have shown that the light "cone" is more appropriate for our purposes. In our case, therefore, instead of rotation we must understand velocity.

At first sight it seems implausible that position should be capable of being confused with velocity in this way and perhaps it will always remain so. However such confusion of qualitatively different observables is already familiar in the theory of relativity where the strict dichotomy of time from space has disappeared. Moreover the interchange of position and velocity goes back to a period before relativity in the shape of the constancy of the sum of potential and kinetic energy. Notwithstanding the implausibility or otherwise of the idea, however, it would seem that it puts in our hands a simple explanation of the recession of the distant nebulae.

The simplest way of understanding this is to take again the extreme case of observers situated at opposite poles of the universe and to appeal to the principle of relativity, that is, that what these observers see of each other should be a function only of their relative situations. The new feature introduced by the knowledge that space is elliptic is that what is locally displacement for one is interpreted locally by the other as velocity and vice-versa. Since they both see the other as at a great distance it would seem that they both see the other also in rapid motion, the motion of the one being the displacement of the other.

The case of observers at intermediate distances is not so straight-forward because distance and velocity are for them not completely reversed. This brings to light certain subtleties in the concepts of distance and velocity when these are large, which could be neglected in the extreme case, and if the simple relativity argument were used confusion would arise. The matter is however made clear by a return to the Clifford concept of parallel translation of the axes of reference. We find there that during a translation the parallel is rotated through an angle proportional to the amount of the translation; in fact, a parallel translation in elliptic space consists of a twist about a screw with pitch unity. Now the rotation part of the twist represents acceleration of the reference frame in our geometry so the proportionality of recession velocity and distance observed in nature follows directly. In particular, of course, the translation which carries an axis into its polar is a twist through a right angle, that is, an acceleration to the maximum velocity, so this argument agrees with the one in the previous paragraph, in the case of polar observers.

5. *Discussion.*—The last three sections contain the main points of the proposed modification of the theoretical basis of relativity which removes the arbitrary features of the older cosmologies and makes the recession of the distant nebulae a purely observational illusion. Although these changes are drastic it has been shown that they do not conflict with accepted principles where these principles can be subjected to experimental test and, furthermore, detailed consideration shows them to be very much in accord with the spirit of special relativity in its earlier form. The special theory emphasized that observation depended on the propagation of light rays and that concepts such as absolute spatial length which could not be measured by sending and receiving light signals should not be represented. In constructing cosmologies which rest on the supposition that space and time are independent variables for distant events even though it is certain that we cannot separate them observationally we were thus abandoning the relativity principle. The present hypothesis of founding the metric in the space of the light "cone" is thus a return to it.

The present paper has deliberately not dealt with more than the very first principles of the new approach. It remains to show that these principles are

not so restrictive as to exclude explanations of other observable phenomena, such as the crowding together of distant matter reported by Ryle. Although this work has already been largely completed an exposition of it requires, as a preliminary, detailed discussions of the fundamentals of geometry and algebra which are probably outside the scope of an astronomical paper and which therefore may have to be published elsewhere. It is hoped, nevertheless, that the further astronomical consequences of the new proposals can be published at an early date.

Army Operational Research Group,
Broadoaks, Parvis Road,
West Byfleet, Surrey:
1956 *October.*

References

(1) Ryle, M., *Observatory*, **75**, 137 (1955).
(2) Cayley, A., " Sixth Memoir upon Quantics ", *Phil. Trans.*, **149**, 61 (1859).
(3) Lie, S., " Theorie der Transformationsgruppen ", Bd. III, Leipzig (1893).
(4) Clifford, W. K., " Preliminary sketch of bi-quaternions ", *Proc. Lond Math. Soc.*, **4**, 381 (1873).

THE SPECTRUM OF QUIESCENT PROMINENCES*

M. A. Ellison and J. H. Reid

Summary

Spectrophotometric results are given for 17 quiescent prominences. The spectra mostly refer to an apparent height of 30 000 km above the solar limb. Equivalent widths have been deduced for the Balmer lines Hα to H_{ϵ} and for the H and K lines of Ca II. The Balmer decrement is discussed.

Instrumental profiles have been determined for the spectrograph in the two orders and with two different widths of slit. The observed prominence profiles in Hα, Hβ and Hγ have thus been corrected for the effects of instrumental broadening.

The corrected mean profiles of Hβ and Hγ give a reasonably close fit for a Maxwellian distribution of velocities of 10 000 deg. K. The Hα profiles, on the other hand, show strong self-absorption effects and give values of the optical depth (C) ranging from 3 to 10. The number of emitting atoms in Hα is of the order 6000 per c.c.

The true half-widths of the H and K lines are considerably less than the resolving power of the present equipment. However, the mean ratio of the equivalent widths, $W'\lambda(\mathrm{H})/W'\lambda(\mathrm{K}) = 0{\cdot}75$, indicates a degree of self-absorption greater than in Hα. For some of the prominences studied the distribution of central intensity in the H and K lines is found to be quite different to that in the hydrogen lines. Differences in composition, or of excitation, within these prominences are suggested.

* The full text of this paper appears in *Publications of the Royal Observatory, Edinburgh*, 2, No. 2, 1957.

ERRATA

***M.N.*, 115,** No. 4, 1955:

A. M. Walker and Andrew Young, *The analysis of the observations of the variation of latitude.*

p. 455, sub-heading to Table I, and line 6 from bottom,

for	$0''{\cdot}01$,
read	$0''{\cdot}001$.

M.N., **116,** No. 2, 1956:

Proceedings of Observatories, *Report of Norman Lockyer Observatory.*

p. 190, line 1,

for	λ Cassiopeiæ,
read	γ Cassiopeiæ.

NOTICE TO AUTHORS

1. *Communications.*—Papers must be communicated to the Society by a Fellow. They should be accompanied by a summary at the *beginning* of the paper conveying briefly the content of the paper, and drawing attention to important new information and to the main conclusions. The summary should be intelligible in itself, without reference to the paper, to a reader with some knowledge of the subject; it should not normally exceed 200 words in length. **Authors are requested to submit MSS. in duplicate. These should be typed using double spacing and leaving a margin of not less than one inch on the left-hand side. Corrections to the MSS. should be made in the text and not in the margin.** Unless a paper reaches the Secretaries more than seven days before a Council meeting it will not normally be considered at that meeting. By Council decision, MSS. of accepted papers are retained by the Society for one year after publication; unless their return is then requested by the author, they are destroyed.

2. *Presentation.*—Authors are allowed considerable latitude, but they are requested to follow the general style and arrangement of *Monthly Notices*. References to literature should be given in the standard form, including a date, for printing either as footnotes or in a numbered list at the end of the paper. Each reference should give the **name and initials** of the author cited, irrespective of the occurrence of the name in the text (some latitude being permissible, however, in the case of an author referring to his own work). The following examples indicate the style of reference appropriate for a paper and a book, respectively :—

A. Corlin, *Zs. f. Astrophys.*, 15, 239, 1938.

H. Jeffreys, *Theory of Probability*, 2nd edn., section 5.45, p. 258, Oxford, 1948.

Alternatively, the Harvard reference system may be used.

3. *Notation.*—For technical astronomical terms, authors should conform closely to the recommendations of Commission 3 of the International Astronomical Union (*Trans. I.A.U.*, Vol. VI, p. 345, 1938). Council has decided to adopt the I.A.U. 3-letter abbreviations for constellations where contraction is desirable (Vol. IV, p. 221, 1932). In general matters, authors should follow the recommendations in *Symbols, Signs and Abbreviations* (London : Royal Society, 1951) except where these conflict with I.A.U. practice.

4. *Diagrams.*—**These should be designed to appear upright on the page, drawn about twice the size required in print and prepared for direct photographic reproduction except for the lettering, which should be inserted in pencil. Legends should be given in the manuscript indicating where in the text the figure should appear.** Blocks are retained by the Society for 10 years; unless the author requires them before the end of this period they are then destroyed.

5. *Tables.*—**These should be arranged so that they can be printed upright on the page.**

6. *Proofs.*—Costs of alteration exceeding 5 per cent of composition must be borne by the author. Fellows are warned that such costs have risen sharply in recent years, and it is in their own and the Society's interests to seek the maximum conciseness and simplification of symbols and equations consistent with clarity.

7. *Revised Manuscripts.*—When papers are submitted in revised form it is especially requested that they be accompanied by the original MSS.

Reading of Papers at Meetings

8. When submitting papers authors are requested to indicate whether they will be willing and able to read the paper at the next or some subsequent meeting, and approximately how long they would like to be allotted for speaking.

9. Postcards giving the programme of each meeting are issued some days before the meeting concerned. Fellows wishing to receive such cards whether for Ordinary Meetings or for the Geophysical Discussions or both should notify the Assistant Secretary.

CONTENTS

PAGE

Meeting of 1956 November 9:

Fellows and Junior members elected 591

Presents announced 591

Meeting of 1956 December 14:

Fellows and Junior member elected 592

Presents announced 592

R. O. Bishop, Chromospheric emission and its dependence on latitude 593

B. E. J. Pagel, A model atmosphere for the solar limb based on continuum observations 608

M. A. Ellison, A study of a blow-off prominence (1956 May 18) 624

J. T. Jefferies, The D_3 emission of prominences 629

D. E. Blackwell and D. W. Dewhirst, An examination of the observational evidence for the accretion theory of the solar corona 637

E. M. Lindsay, A catalogue of stellar-like emission objects in the Small Magellanic Cloud 649

David S. Evans, The sense of rotation of NGC 253 659

W. Rindler, Visual horizons in world-models 662

C. Gilbert, The gravitational field of a star in the expanding universe 678

C. Gilbert, Dirac's cosmology and the general theory of relativity 684

E. R. R. Holmberg, A new explanation of the recession of the nebulae 691

Summary of paper in *Publications of the Royal Observatory, Edinburgh,* **2,** No. 2, 1957: **M. A. Ellison and J. H. Reid,** The spectrum of quiescent prominences 699

Printed in London by Taylor & Francis, Ltd., 1957 May